AD-784 581

REPORT OF TASK FORCE ON TEST AND EVALUATION

Office of the Director of Defense Research
and Engineering
Washington, D. C.

2 April 1974

AD784581

DEFENSE SCIENCE BOARD

REPORT OF TASK FORCE ON TEST & EVALUATION

APRIL 2, 1974

OFFICE OF THE DIRECTOR OF
DEFENSE RESEARCH & ENGINEERING WASHINGTON, D. C.

97

THE DEPUTY SECRETARY OF DEFENSE
WASHINGTON, D. C. 20301

MAR 3 , 1974

MEMORANDUM FOR THE CHAIRMAN, DEFENSE SCIENCE BOARD

THROUGH: DIRECTOR OF DEFENSE RESEARCH AND ENGINEERING

SUBJECT: Report of the Defense Science Board Task Force on Test and Evaluation

I have reviewed the subject report and consider it to be a very worthwhile effort. Implementation of its recommendations concerning (1) reliability planning and testing, (2) orderly and systematic software development and testing, (3) the use of functional specifications wherever possible, and (4) early limited production for operational test and evaluation should produce important benefits in our current efforts to reduce both acquisition and total life cycle costs of DoD systems.

The report will receive widespread distribution in the Department of Defense.

I would like you to express my appreciation to the Chairman and to all of the members of the Task Force for their participation in the preparation of this report, which I know required the contribution of large amounts of their time and capabilities. Their willingness to put their talents at the service of the Government to develop their recommendations to improve the system acquisition process is greatly appreciated.

W. P. Clements Jr.

ia

OFFICE OF THE DIRECTOR OF DEFENSE RESEARCH AND ENGINEERING
WASHINGTON, D. C. 20301

18 March 1974

MEMORANDUM FOR SECRETARY OF DEFENSE

THROUGH: DIRECTOR OF DEFENSE RESEARCH AND ENGINEERING

The attached report of the Defense Science Board Task Force on Test and Evaluation was prepared at the request of the Director of Defense Research and Engineering. The Task Force was chaired by Dr. Eugene G. Fubini and included members of industry, the Services, and the Office of the Deputy Director (Test and Evaluation), ODDR&E.

The Task Force has summarized and delineated procedures to be observed and general guidelines to be followed for the use of members of the Department of Defense and the developers of weapons systems in preparing, reviewing and monitoring the test and evaluation aspects of development programs. A check list of items that must be covered has been prepared as an additional aid.

The Task Force has endorsed the policies of Department of Defense Directive 5000.3, and developed guidelines to be used in conjunction with these policies. The Task Force noted, for example, that programs which preceded publication of DOD Directive 5000.3 sometimes suffered from organizational breaks with the result that information developed in testing did not reach senior Service management levels early enough to head off significant delays and increased costs. The provisions of DOD Directive 5000.3 regarding test reporting procedures, supported by the Task Force guidelines on this subject, should sharply reduce or eliminate this cause of difficulty.

I wish to call your attention to the recommendations concerning a few broad issues that are of particular significance, because they suggest that changes in our present practices are desirable. The issues and recommendations dealing with them are:

(1) Testing the reliability of systems; the Task Force recommends the development of a reliability growth plan as part of system planning. The plan would include the demonstration of achievement of interim reliability goals, (set at a level lower than the ultimate) prior to commencement of limited production; and a subsequent demonstration of achievement of ultimate reliability goals prior to commencement of full production.

(2) Software development and testing; the Task Force recommends that software, like hardware, be developed under an orderly program plan with monitoring by scheduled milestones.

(3) Early limited production; the Task Force recommends early limited production for operational test and evaluation in the many cases where this is possible without very large early commitment of funds.

(4) Writing of specifications; the Task Force recommends that functional specifications be used in place of design specifications whenever that can be done.

If the recommendations of the Task Force on these four issues are followed, important consequences in the budgeting and scheduling of programs will result.

The report has been approved by the Defense Science Board and I recommend it for your consideration.

Solomon J. Buchsbaum
Chairman, Defense Science Board

ic

DEFENSE SCIENCE BOARD

REPORT OF THE TASK FORCE ON TEST AND EVALUATION

OFFICE OF THE DIRECTOR OF DEFENSE RESEARCH AND ENGINEERING

WASHINGTON, D. C.

ii

OFFICE OF THE DIRECTOR OF DEFENSE RESEARCH AND ENGINEERING
WASHINGTON, D. C. 20301

13 February 1974

MEMORANDUM FOR THE CHAIRMAN, DEFENSE SCIENCE BOARD

SUBJECT: Final Report of the Task Force on Test and Evaluation

On November 14, 1972, Dr. Foster asked me to undertake the responsibility of chairing a DSB Task Force aimed at setting some general rules for T&E activity in DoD. Since that time, the Task Force has been organized and 18 weapon systems examined from the point of view of their T&E activities and their effect on the success of the project itself.

Partially from the examination of these projects, but more especially from the experience of its members, the Task Force drew a number of conclusions and guidelines designed for members of T&E organizations who in the future will be charged with the responsibility of formulating, approving and monitoring T&E programs. The Task Force endorses the policies set forth in DoD Directive 5000.3; most of its efforts were devoted to developing guidelines to be used in conjunction with these policies. These guidelines represent a general consensus of the Task Force members but not every member will agree specifically with every item.

The Task Force found it useful to divide the final report into parts: First, a general section that includes nine short sections written in a form of essays and two appendices - also written in the same form. Second, a list of rules which are applicable to most or all weapon systems. This second part is written in the form of a check list where rules are first given and then followed by examples of applications. The Task Force believes that this report will set useful guidelines to insure that T&E programs are properly prepared and avoid many of the errors made in the past.

In addition to this report, nine additional volumes have been prepared by the Task Force not to be used as a DSB report but to be issued by the T&E organization of OSD. These nine volumes deal with specific categories of weapon systems; they are also prepared in check list form with general rules followed by examples.

Since this is the first report of this kind, it is not complete. I would urge that the DDR&E and Service staffs be invited to collect rules similar to those written in this report so that a second edition of these check lists and essays can be prepared in two years. If this procedure is followed, the quality of the report and its supplements will automatically increase. We hope that this first version forms a useful base on which to build future work.

We enjoyed working with General Starbird and his staff and look forward to receiving comments both from the Board and Members of DDR&E who will review it.

G. Fubini

Eugene G. Fubini
Chairman, Task Force on
Test and Evaluation

ACKNOWLEDGMENT

The Chairman wishes to thank the many individuals and their organizations who contributed to this study. This report could not have been prepared without their full cooperation and participation. First, I would like to express my gratitude to the members of the DSB Task Force (see page vii) and their parent organizations, for their invaluable efforts. They contributed a great deal of their own time and talents to this effort and their suggestions and recommendations on how to improve T&E are appreciated.

I would also like to acknowledge the major contributions of the individuals who worked closely with the Task Force members and express thanks to their parent organizations for making it possible for them to serve. They include Mr. Alexander Alexandrovich, Grumman; Dr. Carl Benning, Texas Instruments; Mr. James Bitonti, IBM; Mr. John Buchta, G.E.; Mr. Robert Buzard, LTV; Mr. Victor Friedrich, Office of Assistant Secretary, Army (R&D); Mr. Thomas Glass, Caterpillar Tractor Corp.; Mr. Harold Johnson, Caterpillar Tractor Corp.; Mr. Carrol Killough, Caterpillar Tractor Corp.; Mr. E. W. Neubert, NASA; Mr. Jacob Staab, Caterpillar Tractor Corp.; LCdr. Fred West, CNA.

Mr. Howard Kreiner of ODDR&E and Dr. Joseph Navarro of System Planning Corp. have been not only of tremendous help in the management of this work, but have also contributed many original concepts, precepts and suggestions.

This study could not have been undertaken or conducted without the assistance and cooperation of the many, many others. It would be impossible individually to acknowledge the many people, military and civilian, government and contractor staff, who spent many hours with the DSB Task Force members discussing all aspects of T&E as it relates to the various weapon systems programs. Their information as supplied to the Task Force, is the basic material from which the report has been fashioned.

Finally, I would like to express my warm gratitude for the technical assistance provided by the System Planning Corporation, the administrative assistance provided by ODDR&E(T&E), and the encouragement and support provided by Lt.Gen. Alfred D. Starbird (Ret.), ODDR&E DDT&E.

E. G. Fubini
Chairman

MEMBERSHIP

Task Force on T&E Guidelines

Dr. Eugene G. Fubini, Chairman	President E.G. Fubini Consultants, Limited
Mr. J. Fred Bucy, Jr.	Executive Vice President Texas Instruments Incorporated
Dr. George E. Burks	Vice President (Retired) Caterpillar Tractor Company
Mr. Charles Carnahan	Director of Launch Vehicles Martin Marietta Corporation
Mr. B. O. Evans	President, Systems Development Div. International Business Machines
Mr. John Fosness	President, Columbus Aircraft Div. Rockwell International Corporation
Mr. Charles DiBona	President, Center for Naval Analyses
Mr. Robert Gigliotti	Director, Marine Corps Operations Analysis Group, Center for Naval Analyses
Mr. David R.S. McColl	Deputy Assistant Secretary of the Air Force (R&D)
Mr. Charles L. Poor	Deputy Assistant Secretary of the Army (R&D)
Dr. Peter Waterman	Special Assistant, Office of the Secretary of Navy (R&D)
VADM George Wauchope (Ret.)	Director of Ship Construction and Terminal Design, Farrell Lines
Dr. Lloyd Wilson	Vice President R&D Associates
RADM Forrest S. Petersen	ODDR&E, ODD(T&E), Asst. Dir., SSST&E
Mr. Howard W. Kreiner	ODDR&E, ODD(T&E), Exec. Secretary of the Task Force

CONTENTS

LIST OF RELATED REPORTS

Published by Deputy Director (Test and Evaluation)

T&E GUIDELINES FOR MISSILE WEAPON SYSTEMS

T&E GUIDELINES FOR AIRCRAFT SYSTEMS

T&E GUIDELINES FOR SHIP SYSTEMS

T&E GUIDELINES FOR GROUND VEHICLE SYSTEMS

T&E GUIDELINES FOR ASW SYSTEMS

T&E GUIDELINES FOR AIRBORNE ECM SYSTEMS

T&E GUIDELINES FOR AIRBORNE GENERAL SURVEILLANCE RADAR SYSTEMS

T&E GUIDELINES FOR COMMAND AND CONTROL SYSTEMS

T&E GUIDELINES FOR COMMON TEST GEAR

I. EXECUTIVE SUMMARY

The Defense Science Board Task Force on Test and Evaluation was established at the request of Dr. Foster, Director of Defense Research and Engineering, on behalf of Lieutenant General Alfred D. Starbird (Ret.), Deputy Director (Test and Evaluation) to develop guidance on test and evaluation through examination of a group of representative weapon systems acquisition programs.

The report assumes a significant amount of knowledge on the part of the reader about existing directives and T&E procedures. The emphasis is on listing those T&E items that past experience has indicated had a profound effect on the success of a program.

This report presents guidance on T&E at two distinct levels. At the most general level, this report (Chapter III) discusses a number of issues which are appropriate for <u>all</u> weapon systems acquisition programs, and are generally matters of basic policy. These issues are:

A. Reliability
B. Computer software
C. Human factors
D. The "T&E Gap"
E. Functional specifications versus design specifications
F. Offense/defense testing
G. Portable instrumentation
H. Ship testing
I. Test Planning

Next, a general checklist of items is presented (Chapter IV) which is organized for a rapid overall review of T&E aspects, generally applicable to <u>all</u> systems development and deployment. The T&E expert in reading this chapter will find many precepts which will strike him as being too obvious to be included in checklists of this type. These items are included because many examples were found where even the obvious has been neglected, not because of incompetence or lack of personal dedication by the people in charge of the program, but because of financial and temporal pressures which forced competent managers to compromise on their principles.

A. SYSTEM RELIABILITY

One of the major factors contributing to degraded weapon systems performance is the lack of system reliability, maintainability, and serviceability, three of the major components of availability. The lack of sufficient reliability has been observed in many of the systems studied by the Task Force.

It should be emphasized that the lack of reliability is not measured only by random failures of components but also by the failures induced by poor hardware design, poor software design, operator errors, wear out, and failure to appreciate the severity of environmental conditions. The above failure modes proved difficult and expensive to overcome when they were allowed to persist into the production article.

Ordinarily, reliability specifications are included in the development contract. For some systems, these requirements tend to be far in excess of what is truly needed or achievable in the program. As a result, reliability specifications set by the developing agency were not met, were progressively relaxed, and, in some instances, were never met. As a consequence, realistic reliability goals were not set, and the program lacked a basis for achievement of realistic goals.

The Task Force therefore concludes that the test and evaluation monitors must require that functional (as contrasted with design) reliability goals be defined, in terms of such operational measures as the probability of completing a mission of specified duration, and that testing adequate to demonstrate achievement of these goals be accomplished successfully.

It is not expected that final operational goals will be achieved during the early stages of the R&D program, but it is necessary that the improvement of reliability be planned during the development and engineering phases, be monitored during these phases, and its achievement proven by testing prior to the major production decision.

Reliability is not a uniquely fixed property of a system, but is achieved progressively in the development of a complex system. Consequently, interim goals, and tests based on these interim goals, must be devised to allow tracking of reliability growth through the program. The alternative

of having only a final goal, which is not demonstrable at early stages of the program, allows (if not encourages) contractor and developer alike to overlook the steps necessary before the production decision to achieve the final goal.

Therefore, the progressive attainment of reliability goals must be reviewed at critical points or milestones of the program.

This proposal would, it may be noted, permit the Services to obtain full production approval even prior to the end of the development program, provided reliability growth was tracking well, and thereby reduce the time to operational capability that would have been required if one had to strive for the last most difficult reliability growth.

B. COMPUTER SOFTWARE

Whereas the hardware development was for the most part scheduled, monitored, tested and regularly evaluated, the software development was not.

The Task Force has outlined a software development procedure which should provide for orderly concept program definition, and for continuous testing and monitoring of the software program development, to provide assurance that adequate, efficient, reliable operation will be possible. The increased percentage of development cost introduced by software makes establishment of a suitable procedure a matter of utmost importance.

C. HUMAN FACTORS

User interaction through active participation in the design and execution of test programs is important in all weapon system developments. In systems with a high degree of human interaction--such as Command and Control systems--it is vital that it start with the system design.

D. THE TEST AND EVALUATION GAP

A test and evaluation gap may develop in acquisition programs for expendable equipment between the end of the basic R&D/IOT&E phase and the beginning of the follow-on OT&E, if IOT&E is conducted with R&D prototypes

and no provision is made to obtain production articles until after successful IOT&E is complete. This gap, during which no testing occurs, lasted about 2 years on one recent program. The time lost in maturing the production system and the costs to the contractor and the government from the stopping and starting of hardware construction activities as the program moves from R&D to production are highly undesirable.

E. FUNCTIONAL SPECIFICATIONS VERSUS DESIGN SPECIFICATIONS

Typically, the contractor who is to produce a new system has been given a set of design specifications which the hardware must meet. The contracting service believes that if these design specifications are met, the resulting functional capabilities of the hardware will meet the service needs. Unless the contractor and the government specifically agree otherwise, the government assumes the responsibility of proving that the design specified will perform according to a set of functional specifications (the latter not being binding to the producer).

If the system does not meet functional specifications, the resulting problem can be so serious that one should conclude that the government should never take the responsibility for the assertion that a specific design meets a specific performance.

F. OFFENSE/DEFENSE TESTING

To comply most fully with the spirit of the DoD policy, it would be ideal to have test ranges established with the purpose of maintaining in the field and continuously updating systems based on the most modern technology both for defense and offense. For example, it would be necessary to provide inter-netted defense complexes to test a wide variety of offensive weapons. We would require the test ranges to be capable of testing new defense systems against a similar large variety of offensive devices. The costs of these facilities could be overwhelming and may well be not justifiable in some cases.

G. PORTABLE INSTRUMENTATION

In some cases, in order to have a realistic environment, possibly in simulating a NATO area battle scenario or an amphibious landing, it is necessary to have a portable range instrumentation system available so that the tests can be conducted on and over terrain that provides a realistic operational environment.

H. SHIP TESTING

DoD Directive 5000.3 states that "to the degree practicable first generation subsystems will have been approved for service use prior to the initiation of integrated operational testing." Subsystem approval for service use, by application of other provisions of the Directive, should be preceded by extensive development and operational test and evaluation. The Task Force urges that "first generation" should be liberally interpreted to include subsystems previously approved for service use but which have been "improved" or modified for the new application.

"When combat system complexity warrants, there is to be constructed a combat system test installation wherein the weapon, sensor, and information processing subsystems are integrated through their interfaces in the manner expected in the ship class." The Task Force believes that all combatant classes and most auxiliary classes of ships equipped for ocean use would be of sufficient complexity to warrant such test installation.

The Task Force would add that where possible, in the case of a large number of ships in a class, no more follow-on ships than necessary for economy and early deployment be contracted before completion of this phase of testing.

I. TEST PLANNING AND SCHEDULING

The review of past programs indicated widespread inadequate early planning for test and evaluation.

There are a number of actions that should be taken to improve early planning and test conduct. DoD Directive 5000.3 requires that the DCP prepared at the time of the program initiation ". . . will also provide a summary statement of test objectives, schedules, and milestones."

For this summary to be most meaningful, it is necessary that all agencies who will be involved in the tests be consulted to identify testing time, funds, and resources required for the program.

Many checklist items are contained in this report as reminders of those elements that should be considered in developing this overall plan upon which the program is scheduled and costed. Some of these items cover such things as:

- Ensure that the whole system, including the user people, is tested. Realistically test the complete system, including hardware, software, people and all interfaces. Get user involvement from the start and understand user limitations.
- Ascertain that sufficient time and test articles are planned. When the technology is stressed, the higher risks require more test articles and time.
- In general, parts, subsystems and systems should be proven in that order before incorporating them into the next higher assembly for more complete tests. The instrumentation should be planned to permit diagnosis of troubles.
- Major tests should never be repeated without an analysis of failures and corrective action. Allow for delays of this nature.

It is essential that DSARC actions protect the time and the funds provided for T&E from encroachment due to overruns of time and money in other phases of the program.

II. INTRODUCTION

The Defense Science Board Task Force on Test and Evaluation was established at the request of Dr. Foster, Director of Defense Research and Engineering, on behalf of Lieutenant General Alfred D. Starbird (Ret.), Deputy Director (Test and Evaluation) to develop guidance on test and evaluation through examination of a group of representative weapon systems acquisition programs. (See Appendix A for Terms of Reference.) This report presents the findings of the Task Force through its efforts over a period since 18 December 1972.

The purpose of the report is to offer some guidance to all elements of the Defense Department whose task is to prepare, monitor and execute T&E plans for service use and for presentation to the DSARC.

The report assumes a significant amount of knowledge on the part of the reader about existing directives and T&E procedures. The emphasis is on listing those T&E items that past experience has indicated had a profound effect on the success of a program. Accordingly, it is hoped that these guidelines will be used by the Office of the Secretary of Defense, and the Services, and thus eventually will improve the quality of T&E plans, speed up the approval process of programs and reduce the chances that major difficulties will arise during development programs.

The Task Force found that there was a need for checklists which could be used to assist in the monitoring of the T&E portion of the acquisition program. The guidelines and checklists presented here are the results of lessons hard learned, from examination of weapon systems programs which reflected cost and schedule overruns, inadequate reliability and other defects, as well as those whose histories give examples of methods and procedures for overcoming these problems.

This report presents guidance on T&E at two distinct levels. At the most general level, this report (Chapter II) discusses a number of issues which are appropriate for all weapon systems acquisition programs, and are generally matters of basic policy. The DSB Task Force preferred to present

its content in the form of discussions rather than as a set of checklist items. These issues are:

A. Reliability
B. Computer software
C. Human factors
D. The "T&E Gap"
E. Functional specifications versus design specifications
F. Offense/defense testing
G. Portable instrumentation
H. Ship testing
I. Test planning and scheduling

Next, a general checklist of items is presented (Chapter IV) which is organized for a rapid overall review of T&E aspects, generally applicable to <u>all</u> systems development and deployment. The subjects cover the following areas:

A. General planning
B. Scheduling
C. Criteria
D. Resources
E. Costs
F. Issues
   - Performance
   - Operational Realism
     - General
     - Personnel
     - Threat and environment
G. Reporting

The following brief discussion may help clarify the different emphasis of testing on items as the program develops.

<u>Conceptual Phase Before DSARC I</u>

Tests and plans as the service may feel are necessary to support the DCP, or equivalent documentation, related to the concept definition including both operational and technical aspects and their mutual interaction.

Validation Phase Between DSARC I and DSARC Approval for Full-Scale Engineering Development

Tests and plans related to the validation of the concept. Tests must confirm that the operational concept is sound, that all basic technologies have been validated and that materials, components and subassemblies have been tested to such an extent that the related technical risks are minimized. Plans for tests during the full-scale development should be prepared during this phase.

Between DSARC Approval of Full-Scale Engineering Development and DSARC Approval of Substantial Production/Deployment

Testing of materials, components and subassemblies made on items which are in early production engineering stage or ready for it. In addition, tests must identify engineering problems which appear only when the system is "all up" and investigate the character of these problems; the tests will be followed by demonstrations to confirm the readiness of the items for production. In this phase, the operational character of the tests is paramount and an attempt must be made to identify and investigate the operational problems and to assess the eventual operational suitability and effectiveness of the final product.

Production/Deployment Phase After DSARC Approval of Substantial Production/Deployment

Tests with the same purpose as those of the preceding phase but in this case the articles being tested are the final version of production engineering and demonstration tests of operational capability plan an even more important part. Problems of maintenance, reliability and support will be extremely important as are those associated with organizational and employment concepts.

Not all of the systems rigorously follow the above DSARC cycle. One such example is Command and Control systems. To the extent that this type of system goes through the DSARC procedure it is important to remember that the system has to be evolutionary in nature and flexible to accommodate a wide range of users, and because of this, systems (such as C&C) cannot be tested as a typical weapons system; however, it must always be considered and tested as a total system.

In conclusion, the checklists contained in this report should remind the elements of the Defense Department who prepare and execute the plans or who monitor them of a variety of problems which may appear and call their attention to those problems which have been neglected in the past.

NO ATTEMPT IS MADE TO INCLUDE ALL POSSIBLE PROBLEMS; THE GUIDELINES AND CHECKLISTS ARE BASED ON LESSONS LEARNED FROM PAST EXPERIENCES AND PROBLEMS. THEREFORE, IT IS EXPECTED THAT NEW PROBLEMS WILL APPEAR. However, it is hoped that this report will be a useful tool to focus the attention of the reader not only on old problems but also on the new ones.

III. GENERAL ISSUES

A. SYSTEM RELIABILITY

One of the major factors contributing to degraded weapon systems performance is the lack of system reliability, maintainability, and serviceability, three of the major components of availability. The lack of sufficient reliability has been observed in many of the systems studied by the Task Force.

It should be emphasized that the lack of reliability is not measured only by random failures of components but also by the failures induced by poor hardware design, poor software design, operator errors, wear out, and failure to appreciate the severity of environmental conditions. The above failure modes proved difficult and expensive to overcome when they were allowed to persist into the production article.

Ordinarily, reliability specifications are included in the development contract. For some systems, these requirements tend to be far in excess of what is truly needed or achievable in the program. As a result, reliability specifications set by the developing agency were not met, were progressively relaxed, and, in some instances, were never met. As a consequence, realistic reliability goals were not set, and the program lacked a basis for achievement of realistic goals.

The Task Force therefore concludes that the test and evaluation monitors must require that functional (as contrasted with design) reliability goals be defined, in terms of such operational measures as the probability of completing a mission of specified duration, and that testing adequate to demonstrate achievement of these goals be accomplished successfully.

It is not expected that final operational goals will be achieved during the early stages of the R&D program, but it is necessary that the improvement of reliability be planned during the development and engineering phases, be monitored during these phases, and its achievement proven by testing prior to the major production decision.

Reliability is not a uniquely fixed property of a system, but is achieved progressively in the development of a complex system. Consequently, interim goals, and tests based on these interim goals, must be devised to allow tracking of reliability growth through the program. The alternative of having only a final goal, which is not demonstrable at early stages of the program, allows (if not encourages) contractor and developer alike to overlook the steps necessary before the production decision to achieve the final goal.

Therefore, the progressive attainment of reliability goals must be reviewed at critical points or milestones of the program. Specifically, these are:

1. At the time the service requests initiation of engineering development, it should be prepared to show a reliability growth plan with sufficient test time and funds to achieve the program goal for reliability achievements.
2. At the time the service requests initiation of limited production, it should be prepared to show:
   (a) By demonstration test results, the system has achieved, at a reasonable confidence level, some percent of the reliability goals for the program, where both confidence level and percent achievement are appropriate to the program.
   (b) There still remains between this time and the end of the development program, sufficient system testing to carry on reliability growth from the point achieved to the program goal for reliability achievement.
3. At the time that the service requests authorization for full-scale production, it should be prepared to show:
   (a) By demonstration test results, the system has achieved, at a reasonable confidence level, nearly all the program reliability goals, if not the final value.
   (b) There still remains between this time and the end of the development program, sufficient system testing to carry on reliability growth from the point achieved, to the program goal for reliability attainment.
   (c) A management plan, test plan and funds to utilize the remaining test time for a vigorous program of reliability growth.

This proposal would, it may be noted, permit the Services to obtain full production approval even prior to the end of the development program. provided reliability growth was tracking well, and thereby reduce the time

to operational capability that would have been required if one had to strive for the last most difficult reliability growth.

If the above recommendations are followed, the percentage of R&D funds required will be higher; however, the total program costs should be lower because of the resulting improved reliability and the associated decreased potential for cost overruns.

B. COMPUTER SOFTWARE

Although most of the programs examined by the Task Force did not use large computer programs, those that did displayed a serious difference in attitude between the development of the computer software and the development of the hardware. Whereas the hardware development was for the most part scheduled, monitored, tested and regularly evaluated, the software development was not. One should not assume that software testing plans are right.

It is more difficult to determine the status of completion of various phases of the software program (as compared to hardware programs), so it is important to explore how the software program is developed and managed as well as how it is being tested. No standard procedure seems to be available within OSD for orderly testing of software items; the Task Force considers this situation unacceptable. Accordingly, the Task Force therefore has outlined a software development procedure which should provide for orderly concept program definition, and for continuous testing and monitoring of the software program development, to provide assurance that adequate, efficient, reliable operation will be possible.

The increased percentage of development cost introduced by software makes the establishment of a suitable procedure a matter of utmost importance. For this reason the procedure suggested is given in this report in some detail, in Annex A. The reader is urged by the Task Force not to assume that the editorial decision of including the procedure in an annex rather than in the text, indicates a low priority for this recommendation.

C. HUMAN FACTORS

The Task Force turned up a surprisingly large number of instances in which designs lacked adequate human factor considerations, and, notable from a T&E point of view, many in which development engineering testing did not lead to early awareness of the problem. The problems were varied: excessive sound levels, insufficient space, or inconvenient access, even poor placement of controls and readouts sufficient to double the manpower requirements for operation of a system.

The solution is obvious: first, more attention should be given to human factors in the initial design, during modifications and updating of equipment; and second, T&E should be planned and conducted so as to ensure that human factor requirements have been adequately considered during design, demonstrated at the first mockup of the system, and monitored throughout subsequent testing.

User interaction through active participation in the design and execution of test programs is important in all weapon system developments. In systems with a high degree of human interaction--such as Command and Control systems--it is vital that it start with the system design.

D. THE TEST AND EVALUATION GAP

A test and evaluation gap may develop in acquisition programs for expendable equipment between the end of the basic R&D/IOT&E phase and the beginning of the follow-on OT&E, if IOT&E is conducted with R&D prototypes and no provision is made to obtain production articles until after successful IOT&E is complete. This gap, during which no testing occurs, lasted about 2 years on one recent program. The time lost in maturing the production system and the costs to the contractor and the government from the stopping and starting of hardware construction activities as the program moves from R&D to production are highly undesirable.

There are three basic alternatives for addressing the acquisition of expendable equipment for the later OT&E phases:

1. Plan at the start of engineering development for additional R&D hardware, to be R&D funded and built for IOT&E and for an additional phase of testing to cover the T&E gap.

Paragraph 5 of DoD Directive 5000.3 recognizes that additional phases of OT&E may be needed prior to the availability of production hardware. In this case, every effort would be made to production tool each subsystem as soon as it could be qualified. In this way, the R&D would gradually evolve into the production configuration.

2. Plan the development and OT&E phases so that DT&E and IOT&E hardware is funded with R&D. Early in the DT&E effort, defend long lead time production funding and seek production funds for low rate pilot production. Again, emphasize early conversion to production configuration so that the evolving production configuration hardware will be available to continue the OT&E immediately after the IOT&E. The testing would be continuous and at a point where all the qualified subsystems were in production, the follow-on OT&E would be initiated.

3. Simply allow the gap to exist, which may be preferred when the effort to reduce the gap would require the commitment to a very large percentage (or amount) of the expected program cost before T&E assurance of a successful product could be obtained.

For further discussion on the T&E gap solutions, the reader is referred to Annex B.

E. FUNCTIONAL SPECIFICATIONS VERSUS DESIGN SPECIFICATIONS

Typically, the contractor who is to produce a new system has been given a set of design specifications which the hardware must meet. The contracting service believes that if these design specifications are met, the resulting functional capabilities of the hardware will meet the service needs. Unless the contractor and the government specifically agree otherwise, the government assumes the responsibility of proving that the design specified will perform according to a set of functional specifications (the latter not being binding to the producer).

If the system does not meet functional specifications, the resulting problem can be so serious that one should conclude that the government should never take the responsibility to tie a design to a performance. An alternative solution is to assign contracts of a system or device on the basis of "form, fit, function and interfaces." Then the interchangeability and performance are clearly the responsibility of the producer. This leads to the following:

GUIDELINE:

<u>When the designer and producer are different, tests should be conducted to ensure that the producer meets the design specifications. The test plan should make provisions for the case when the design specifications are met but the performance is below requirements. In this case it may be necessary to do additional R&D work. Normally, the producer will be assigned this task</u>.

F. OFFENSE/DEFENSE TESTING

The Department of Defense Directive 5000.3 states, "OT&E is that test and evaluation conducted to estimate the prospective system's military utility, operational effectiveness, and operational suitability.... OT&E will be continued as necessary during and after the production period to refine these estimates, to evaluate changes, and to reevaluate the system to ensure that it continues to meet operational needs and retains its effectiveness in a new environment or against a new threat."

Some new systems go through the OT&E without being exposed to <u>any</u> offense/defense environment.

To comply most fully with the spirit of the DoD policy, it would be ideal to have test ranges established with the purpose of maintaining in the field and continuously updating systems based on the most modern technology both for defense and offense. For example, it would be necessary to provide inter-netted defense complexes to test a wide variety of offensive weapons. We would require the test ranges to be capable of testing new defensive systems against a similar large variety of offensive devices.

The costs of these facilities could be overwhelming and may well be not justifiable in some cases. Criterion C-5 in our general checklist refers to this issue and gives the basis for analyses of this tradeoff. An example where this type of activity was in fact conducted and the cost justified was in the test range designed to validate our ABM concepts.

G. PORTABLE INSTRUMENTATION

DoD Directives 5000.1 and 5000.3 stress that OT&E will be conducted in as realistic an operational environment as possible. Although there are a

number of national test ranges available, it is not clear that they could adequately accommodate all new system OT&E programs. In some cases, in order to have a realistic environment, possibly in simulating a NATO area battle scenario or an amphibious landing, it is necessary to have a portable range instrumentation system available so that the tests can be conducted on and over terrain that provides a realistic operational environment.

For these reasons the DSB Task Force recommends serious consideration of such instrumentation. Further, because of the "free play" type testing usually conducted during OT&E, the portable instrumentation must be capable of covering large areas and providing data on player location and events. Such portable instrumentation is especially pertinent to missile and aircraft testing.

H. SHIP TESTING

The testing of ships considered as a system rather than an aggregate of items is a new concept. There could be a tendency not to give serious consideration to Directive 5000.3 because of the many loopholes left in the directive. The Task Force believes that it must restate, at greater length, the procedures given in Directive 5000.3 for testing ships, and emphasize the importance of not bypassing any of the steps.

DoD Directive 5000.3 states that "to the degree practicable first generation subsystems will have been approved for service use prior to the initiation of integrated operational testing." Subsystem approval for service use, by application of other provisions of the Directive, should be preceded by extensive development and operational test and evaluation. The Task Force urges that "first generation" should be liberally interpreted to include subsystems previously approved for service use but which have been "improved" or modified for the new application. It is essential that the DCP for the ship program identify, and make provision for resolution of any remaining uncertainties about the qualification of critical subsystems for inclusion in the ship. Note that the provision of the Directive relates to initiation of integrated testing, rather than to initiation of construction of the lead ship. It is assumed that the lead ship could be well into construction before all equipments were service approved.

"When combat system complexity warrants, there is to be constructed a combat system test installation wherein the weapon, sensor, and information processing subsystems are integrated through their interfaces in the manner expected in the ship class." The Task Force believes that all combatant classes and most auxiliary classes of ships equipped for ocean use would be of sufficient complexity to warrant such test installation.

The foregoing words allow either a land-based or at-sea installation, and, possibly, a good deal of latitude about the detail to be incorporated in the installation. It is recommended that the installation, at a minimum, include accurate, geometrically identical spacing and placement of all critical equipments, at least mockups of other installed equipment in spaces, cable and utility conduits and piping identical to that to be installed in the production ship, antennas, lighting and ventilation as in the production ships (even if augmented for non-test repair and modification), and provision for feeding the test system either real or simulated input as it would occur in operational situations. Real inputs should be used if at all possible and simulated inputs permitted only in cases such as sonar on a land-based test installation.

If the new class of ships incorporates advancements in propulsion technology, there should be a propulsion test site. The Task Force feels that its interpretation of the policy of 5000.3 as it applies to a combat systems test site is equally applicable to a propulsion test site if one is required.

The Directive also states, "for all new ship classes continuing phases of OT&E on the lead ship will be conducted at sea as early in the acquisition process as possible for specified systems or equipments and, if required, full ship operational evaluation to the degree feasible." The Task Force would add that where possible, in the case of a large number of ships in a class, no more follow ships than necessary for economy and early deployment be contracted before completion of this phase of testing. The Task Force also urges that contract methods be devised to minimize the cost impact of changes found necessary in such operational testing.

The Task Force concurs that there should be prototyping of the entire ship and combat system if the new ship's hull design will contain

technological advancements and/or significant scale-ups of previously proven technologies, with operational tests at sea prior to production commitments to follow ships.

I. TEST PLANNING AND SCHEDULING

The review of past programs indicated widespread inadequate early planning for test and evaluation.

The original program estimates were based on incomplete considerations of time and cost implications of the test program. Once the test program requirements were established, there was a great reluctance to modify the schedule or cost estimates. In most cases, the result was inefficient testing and evaluation and cost and schedule overruns.

There are a number of actions that should be taken to improve early planning and test conduct. DoD Directive 5000.3 requires that the DCP prepared át the time of the program initiation ". . . will also provide a summary statement of test objectives, schedules, and milestones."

For this summary to be most meaningful, it is necessary that all agencies who will be involved in the tests be consulted to identify testing time, funds, and resources required for the program.

Many checklist items are contained in later chapters of this report as reminders of those elements that should be considered in developing this overall plan upon which the program is scheduled and costed. Some of these items cover such things as:

- Ensure that the whole system, including the user people, is tested. Realistically test the complete system, including hardware, software, people and all interfaces. Get user involvement from the start and understand user limitations.
- Ascertain that sufficient time and test articles are planned. When the technology is stressed, the higher risks require more test articles and time.
- In general, parts, subsystems and systems should be proven in that order before incorporating them into the next higher assembly for more complete tests. The instrumentation should be planned to permit diagnosis of troubles.
- Major tests should never be repeated without an analysis of failures and corrective action. Allow for delays of this nature.

It is essential that DSARC actions protect the time and the funds provided for T&E from encroachment due to overruns of time and money in other phases of the program.

The DSARC procedures and attitudes can be used in a positive fashion to improve the test planning and scheduling performance by being aware of the situation as discussed above and mainly by insisting upon adequate contingency planning in preparation of the initial DCP, by encouraging thorough updating of the test planning in the Validation Phase before the initiation of full-scale development, and by carefully avoiding the establishment of any deadline or reviews that foster a feeling that testing must be completed by a given date.

IV. GENERAL CHECKLIST ITEMS

The set of checklist items presented in this chapter is oriented toward good procedures and practices relative to T&E. This checklist contains items which for the most part cut broadly across both weapon system types and time phasing of testing. It should serve as a basis for a rapid, if not exhaustive, overall review of a test plan. The organization has been chosen to facilitate just such a quick review, with the expectation that this will be followed by a more thorough examination against the specific checklists. Several of the items in the General Checklist have applicability under several headings (e.g., Scheduling and Resources) and are repeated under each, perhaps with different emphasis. The subjects touched on by the checklist are:

A. General Planning
B. Scheduling
C. Criteria
D. Resources
E. Costs
F. Issues
   - Performance
   - Operational Realism
     - General
     - Personnel
     - Threat and Environment
G. Reporting

NOTE

The T&E expert in reading this chapter will find many precepts which will strike him as being too obvious to be included in checklists of this type. These items are included because many examples were found where even the obvious has been neglected, not because of incompetence or lack of personal dedication by the people in charge of the program, but because of financial and temporal pressures which forced competent managers to compromise on their principles. It is hoped that the inclusion of the obvious will prevent repetition of the serious errors which have been made in the past when such political, economic and temporal pressures have forced project managers to depart from the rules of sound engineering practices.

It is the conviction of the Task Force that, in the long run, taking short cuts during T&E to save time and money will result in significant increases in the over-all costs of the programs and in the delay of the delivery of the corresponding weapon systems to the combatant forces.

A. GENERAL PLANNING

The following are checklist items contained in this section:

1. Effects of Test Requirements on System Acquisition Strategy
2. Test Plan Coverage
3. Test Requirements and Restrictions
4. Trouble Indicators
5. Effect of Incentives on Test and Evaluation
6. Software Testing
7. Requirement for Test Rehearsals

Preceding page blank

1. EFFECTS OF TEST REQUIREMENTS ON SYSTEM ACQUISITION STRATEGY

The acquisition strategy for the system should:

(a) Allow for a sufficient time between the planned end of demonstration testing and major procurement as contracted with limited production decisions so that there is a flexibility for modification of plans which will be required during the test phases of the program;

(b) Ensure that sufficient dollars are available not only to conduct the planned T&E but to allow for the additional T&E which is always required due to failures, design changes, etc.;

(c) Be evaluated relative to constraints imposed by:

- The level of system testing at various stages of the RDT&E cycle,
- The number of test items available and the schedule interface with other systems needed in the tests, such as aircraft, electronics, etc.,
- Support required to assist in the preparation, conduct of the tests, and the analysis of test results;

(d) Be evaluated to minimize the so-called T&E gap caused by a lack of hardware. Specifically, a test gap can result if funds are not applied until the results of IOT&E are known because of the required lead time for production planning, production facilities, and tool and production hardware. (See the T&E gap discussion in Volume 1, Chapter II.)

2. TEST PLAN COVERAGE

Every test plan should include clear statements of:

- The overall purpose of the test
- Critical issues with respect to operational requirements
- The major test objectives
- The schedule of test milestones
- The major resources required
  - Test environment, facilities, and instrumentation
  - Operational environment
- The organizations which will conduct the test program
- The analysis and evaluation approach

3. TEST REQUIREMENTS AND RESTRICTIONS

Tests should:

- Have specific objectives
- List in advance actions to be taken as a consequence of the test results
- Be instrumented to permit diagnosis of the causes of lack of performance including:
  - "Random" failures
  - Design induced failures
  - Wear out failures
  - Operator error failures
  - And those as a result of accidental environmental conditions.
- Not be repeated if failures occur, without a detailed analysis of the failure. Most likely, the failure will not go away. Note that this rule, essential as it is, can be violated if the failure mode analysis reveals that, even if the same failure reoccurs, very useful results can still be obtained about the performance of other subsystems or components.

4. TROUBLE INDICATORS

Establish an early detection scheme for top government and contractor management to determine that a program may be becoming ill.

At this time there may be a good possibility of recovery. Some of the indications of trouble are:

- A test failure
- Any repetitive failure
- A revision of schedule or incremental funding that exceeds the original plan. Predicted downstream recovery may not have a realistic basis.
- Any relaxation of basic requirements such as lower performance, etc.

5. EFFECT OF INCENTIVES ON TEST AND EVALUATION

Improper incentives can warp the proper conduct of the test and evaluation.

In demonstrations, the success criteria should be broader than simply

hit or miss in a single given scenario. Otherwise, the entire program may be skewed to meet the requirements of the selected scenario to the detriment of testing a wider area of the performance envelope.

6. SOFTWARE TESTING

Test and evaluation should ensure that software products are tested appropriately during each phase.

Software has often been developed more as an add-on than as an integral part of the overall system. Software requirements need the same consideration as hardware requirements in the Validation Phase. Usual practices often do not sufficiently provide for testing the software subsystem concept. Facilities available to contractors for software development and verification are becoming increasingly critical to schedule and cost. Note that this topic is treated at greater length in Chapter II and in Annex A.

7. REQUIREMENT FOR TEST REHEARSALS

Test rehearsals should be conducted for each new phase of testing.

The purpose is to shake down the test plan, the instrumentation concept, and the data analysis plan. A secondary, but vital, purpose should be to provide training for the test participants. The pilot run should be scheduled and conducted in such a way that sufficient time will be available to make the necessary changes to the test as dictated by the results of the pilot run.

In the pilot run, particular attention must be given to the range safety aspects so that range safety officials do not destroy a good test because of previously undiscovered momentary deficiencies which might occur during the surveillance of the test article.

Simulation and other laboratory or ground testing should be conducted to predict specific test outcomes. The test run should of course be run to verify the test objectives. Evaluation of the simulation versus the actual test results will help to refine the understanding of the system.

B. SCHEDULING

The following are checklist items contained in this section:

1. Building Block Test Scheduling
2. Component and Subsystem Test Plans
3. Phasing of DT&E and IOT&E
4. Use of Functional Milestones
5. Test Schedule Constraints
6. Requirements for Military Construction Program Facilities
7. Scheduling of Tests Using Government Furnished Equipment
8. Scheduling IOT&E to Include System Interfaces with Other Systems

Preceding page blank

1. BUILDING BLOCK TEST SCHEDULING

The design of the set of tests to demonstrate feasibility prior to DSARC II should be based on a building block concept.

High technical risk items should be tested early and subsequent tests should incorporate more and more of the hardware until the complete system concept has been demonstrated as feasible.

2. COMPONENT AND SUBSYSTEM TEST PLANS

Assure a viable component and subsystem test plan.

Studies show that almost all component failures will be the kind that cannot easily be detected or prevented in full system testing. All experience indicates that new systems will exhibit the "new system syndrome" and that the best return on test investment will come from applying substantial attention to component and subsystem level test effort. Detecting a subsystem or component failure only at the operational test level puts the cost of correcting such failures at the high end of an exponential cost curve.

3 PHASING OF DT&E AND IOT&E

In evaluating test plans, look favorably on phasing where the IOT&E is run in parallel with continued DT&E.

Problems that become apparent in the operational testing can often be evaluated much more quickly and more completely with the instrumented DT&E hardware. This is more attractive where the DT&E is performed with non-expendable hardware like airplanes.

In general, DT and OT plans and schedules must be rejected if they do not make provisions for the occurrence of failures. Plans should include time and money necessary for investigating test failures and making provisions for elimination of the cause before other similar tests take place. (However, see A-3.) Further, it is imperative that a percentage of the total tests (sorties, runs, trials, experiments) be allowed for retesting over and above the number required to successfully complete the program.

 Preceding page blank

This percentage must be related to the probability of achieving success as opposed to failure.

4. USE OF FUNCTIONAL MILESTONES

System milestones should be flexible with respect to time.

In evaluating the adequacy of the scheduling as given by test plans, it is important that milestones be tied to the major events of the weapon system (meeting stated requirements) and not the calendar. The acquisition process should be based on the achievement of major milestones and sufficient time and resources allowed between these milestones. This flexibility must not be hampered by the contracting mechanism. Contractors should be required to demonstrate successful accomplishment of technical milestones before proceeding to the next phase of development.

5. TEST SCHEDULE CONSTRAINTS

The test schedule for the system should:

(a) Allow for a sufficient time between the planned end of demonstration testing and major procurement decisions so that there is a flexibility for modification of plans which may be required during the test phases of the program;

(b) Be evaluated relative to constraints imposed by:

- The number of test items available and the schedule interface with other systems needed in the tests, such as aircraft, electronics, missiles, etc.
- Support required to assist in the preparation, conduct of the tests and the analysis of test results;

(c) As stated previously in A-1, be adjusted to minimize the so-called T&E gap caused by a lack of hardware. Specifically, a test gap can result if funds are not applied until the results of IOT&E are known because of the required lead time for production planning, production facilities, and tool and production hardware.

6. REQUIREMENTS FOR MILITARY CONSTRUCTION PROGRAM FACILITIES

Some systems cannot be fully tested without Military Construction Program (MCP) facilities.

The long lead times to obtain authorization, appropriations, and to construct facilities can pace a program. Many steps and considerable time may be involved in getting facilities ready and test gear in place to start system tests.

If completion of DT&E and the operational testing requires the MCP facility, these matters must be considered in preparing and evaluating the test plan.

7. SCHEDULING OF TESTS USING GOVERNMENT FURNISHED EQUIPMENT

If there are GFE and other government commitments in the proposed contract, be concerned about the following:

(a) Can the gear with required performance be available when required?

(b) Can government supported facilities provide the assistance required at the time needed? If not, is it reasonable to construct the required facilities (test range, instrumentation, building, etc.)? If not, what alternatives are available?

(c) Avoid contract terms on fixed price contracts that vaguely commit the government. Do not include "Government support as required" or "test facilities will be made available when needed."

8. SCHEDULING IOT&E TO INCLUDE SYSTEM INTERFACES WITH OTHER SYSTEMS

Whenever possible, the IOT&E (as well as the FOT&E) of a weapon system should be planned to include other systems which must have a technical interface with the new system.

Thus missiles should be tested on most of the platforms for which they are programmed. Interfaces between systems should receive special attention.

C. CRITERIA

The following are checklist items contained in this section:

1. Criteria for Critical Issues
2. Criteria for Competitive Testing
3. Criteria for Performance Demonstrations
4. Reliability Determinations in IOT&E
5. Expected Value of Testing

1. CRITERIA FOR CRITICAL ISSUES

<u>In evaluating the initial DCP (or its equivalent documentation such as PMs), it is important to ensure that the tests to be conducted during the period from DSARC I to DSARC II address the major critical issues, especially those technological issues which are identified in the DCP</u>.

By the end of the systems Definition Phase, test and evaluation should make certain that "test criteria" are established so there is no question as to what constitutes a test and what performance is to be attained. Each test should have a single objective if possible, and the objective should be simply stated. A plan for the conduct of the test and the data collection, reduction, and analyses must be in sufficient detail that one can readily evaluate the performance of the system and whether or not the test objective can be met. A relationship between the identified performance parameters and the test results should be established prior to the conduct of the test. Further, the set of objectives for each of the tests should be clearly related to the program objective as defined in the DCP. When this relationship is not clear, amplifying data should be required.

2. CRITERIA FOR COMPETITIVE TESTING

<u>When competitive designs are under consideration, criteria for selection should be specified in advance, with critical issues identified for each design</u>.

The DCP, or equivalent documentation, should include the evaluation criteria to be used for the selection of the final system design. They should be based on performance factors which are measurable through testing. A data collection and evaluation plan should be developed which will permit description of the range of acceptable performance for each factor.

3. CRITERIA FOR PERFORMANCE DEMONSTRATIONS

<u>In designing contractually required demonstration tests upon whose outcome may depend large incentive payments, or even program continuation, it is essential to specify broader success criteria than simply hit or miss in a single given scenario.</u>

Preceding page blank

If this is not done, the entire program may be skewed to meet the requirements of the selected scenario, to the detriment of exploring a wider area of the performance envelope. With too much weight attached to a go/no go outcome, temporary hardware, not designed for the final purpose, may be retained beyond the early stages of the program to enhance the probability of successful demonstration.

Demonstrations should be designed to measure overall performance, with statistical weighting to compensate for reduced probabilities of occurrences at edge values of condition parameters.

Contract requirements and incentives should not be weighted heavily on performance at extreme corners of the theoretical performance envelope unless there is a very high payoff for such performance, since excessive effort may be spent on obtaining it.

4. RELIABILITY DETERMINATIONS IN IOT&E

IOT&E can provide valuable data on the operational reliability of weapon systems which cannot be obtained through DT&E.

Apparent operator error failures and apparent random failures should be looked for in the operational tests and investigated to determine if serious problems are underlying reasons for the failures. Especially important is the procedure used to evaluate the operational reliability of the system as determined by the relatively small but significant amount of data obtained through IOT&E and the larger amounts of data on hardware design reliability collected through DT&E. Further, maintenance practices should be carefully studied to assess their impact on the observed operational reliability obtained through IOT&E.

Validation of system life cycle cost should be a primary objective of IOT&E. Inasmuch as procurement cost of any system is only the tip of the iceberg, other costs such as operation and maintenance will, over the life cycle, make up a larger portion of the cost to the taxpayer. Where inordinate expenditures for replacement of high-cost components, heavy operator manning requirements, or high maintenance man-hours per operating hour can

be identified or forecast through IOT&E, this should be done. Where possible, such predictions should be made in quantitative terms.

5. EXPECTED VALUE OF TESTING

Operational testing is essential, but it is also expensive and time consuming.

Be sure in advance that the value received is worth its weight in not-delivered systems. Think in terms of:

(a) Involving operational groups in test planning and in establishing measures of effectiveness, so that the outcome of the tests will be accepted as being operationally significant.

(b) Determining whether the scope of the planned tests will provide sufficient data to justify any change at all in the eyes of potential users.

(c) Comparing the scope of proposed tests against checklists of issues frequently raised at major decision milestones, to assure that the data needed for such decisions will be forthcoming to the extent this is possible from testing alone.

(d) Recognizing in the formulation of test plans that major system decisions are judgments based on a wide range of qualitative considerations, rather than on statistical compilations, and that the outcome and limitations of operational tests must be comprehensive and meaningful to the decision makers as well as to the testing community.

D. RESOURCES

The following are checklist items contained in this section:

1. Identification of Test Resources and Instrumentation Requirements
2. Requirements for Joint Service OT&E
3. Military Construction Program Facilities
4. Government Furnished Equipment
5. Instrumentation Packages for OT&E
6. Test Sample Sizes

Preceding page blank

1. IDENTIFICATION OF TEST RESOURCES AND INSTRUMENTATION REQUIREMENTS

Before DSARC II the test facilities and instrumentation requirements to conduct operational tests should be identified, along with a tentative schedule of test activities.

The applicability of existing test ranges and the adequacy of current facilities and instrumentation should be verified. Insofar as possible, alternative approaches (different ranges, etc.) and instrumentation improvements needed should be specified. Of prime importance are the constraints to be placed on the test because of the range and instrumentation. If range and instrumentation factors are found to cast significant doubt on the meaningfulness of the test data because of a lack of operational realism, the steps necessary to assure meaningful data should be identified and planned prior to DSARC II.

2. REQUIREMENTS FOR JOINT SERVICE OT&E

Joint service operational test and evaluation should be considered for those weapon systems which require new operational concepts involving other services.

Emphasis in the joint tests should include investigations of the impact on the effectiveness of the weapon system of such aspects as CCC, target acquisition, damage assessment, and countermeasures. If joint testing is recommended, an analysis of the impact of this type of demonstration on time and resources needed in the program and the additional resources needed to execute the joint tests should be conducted before DSARC II.

3. MILITARY CONSTRUCTION PROGRAM FACILITIES

Some systems cannot be fully tested without Military Construction Program (MCP) facilities.

As stated before, the long lead times to obtain authorization, appropriations, and to construct facilities can pace a program. Many steps and considerable time may be involved in getting facilities ready and test gear in place to start system tests.

If completion of DT&E and the operational testing requires the MCP facility, these matters must be considered in preparing and evaluating the test plan.

4. GOVERNMENT FURNISHED EQUIPMENT

If there are GFE and other government commitments in the proposed contract, be concerned about the following:

(a) Can the gear with required performance be available when required?

(b) Can government-supported facilities provide the assistance required at the time needed? If not, is it reasonable to construct the required facilities (test range, instrumentation, builting, etc.)? If not, what alternatives are available?

(c) Avoid contract terms on fixed price contracts that vaguely commit the government. Do not include "government support as required" or "test facilities will be made available when needed."

5. INSTRUMENTATION PACKAGES FOR OT&E

The manner in which T&E instrumentation is used can be extremely important in determining the realism possible in the OT&E phases.

The instrumentation package should be fixed early in the design phase of the development; it is difficult and costly to change thereafter. For this reason, instrumentation requirements must be specified early in the program and operational factors must be incorporated early.

6. TEST SAMPLE SIZES

The primary basis for the test sample size is usually based on one or more of the following:

- Analysis of test objectives
- Statistical significance of test results at some specified confidence level.
- Availability of test vehicles, items, etc.
- Support resources or facilities available
- Time available for the test program.

One should not hesitate to terminate a test prior to its completion when it becomes clear that the main objective of the test is unachievable (because of hardware failures, unavailability of resources, etc.), or that additional samples will not change the outcome and conclusions of the test.

E. COSTS

The following are checklist items contained in this section:

1. Budgeting for Test
2. Funds for Correction of Faults Found in Testing
3. Component and Subsystem Test Plans

1. BUDGETING FOR TEST

The DCP and later budgeting documents should be regularly reviewed to ensure that there are adequate identified funds for testing, relative to development and fabrication funds.

A review of previous programs shows that testing funds and test articles have been postponed or eliminated to keep program costs in line as projected development requirements or costs have increased.

2. FUNDS FOR CORRECTION OF FAULTS FOUND IN TESTING

The DCP and later budgeting documents need careful scrutiny to ensure that there are adequate contingency funds to cover correction of difficulties at a level which matches the Industry/Government experience on such contracts. (Testing for difficulty without sufficient funding for proper correction results in band aid approaches which ultimately require correction at a later and more expensive time period.)

Discussions with industry representatives indicate almost universally an erosion process of contingency funds throughout the bidding and negotiation process. This fact has led to enormous financial difficulties to the contractors in "package procurement programs." Today there is a trend toward funding difficulties on Cost Reimbursement Contracts because contractors have been encouraged to be optimistic because of their low legal liability. Further, inadequate contingency funding is being carried by the government.

3. COMPONENT AND SUBSYSTEM TEST PLANS

Assure a viable component and subsystem test plan.

As previously stated, studies show that almost all component failures will be the kind that cannot easily be detected or prevented in full system testing. All experience indicates that new systems will exhibit the "new system syndrome," and that the best return on test investment will come from applying substantial attention to component and subsystem level test effort. Detecting a subsystem or component failure only at the operational test level puts the cost of correcting such failures at the high end of an exponential cost curve.

F. ISSUES: Performance

1. Necessity for Ranges of Criteria
2. Effects of Incentives on Test and Evaluation
3. High Technical Risk Development
4. Proof of Performance on Major Critical Issues
5. Proof of Performance of Software
6. Proof of Performance of Human Factors Concepts

1. NECESSITY FOR RANGES OF CRITERIA

Analytic and empirical studies should be conducted prior to DSARC I to ensure that the range of critical performance characteristics has been specified.

Each performance characteristic so specified should be measurable through bench and laboratory or proving ground testing. The test design and the number of tests should be adequate to provide results with confidence limits compatible with the statements of desired characteristics. Testing in advanced development should be planned to explore performance characteristics over a broad range of environments so as to provide insight into system performance over the expected operational range and not just at a single point.

2. EFFECTS OF INCENTIVES ON TEST AND EVALUATION

Improper incentives can warp the proper conduct of testing and evaluation.

In reviewing contractually required demonstration tests upon whose outcome may depend large incentive payments, or even program continuation, it is essential to specify broader success criteria than simply success or failure in a single given scenario. If this is not done, the entire program may be skewed to meet the requirements of the selected scenario, to the detriment of exploring a wider area of the performance envelope. At the same time, contract requirements and incentives should not be based upon performance at extreme corners of the theoretical performance envelope unless there is a very high payoff of such performance since excessive effort may be spent in obtaining it.

3. HIGH TECHNICAL RISK DEVELOPMENT

When high technical risk is present, development should be structured around the use of prototypes designed to prove the system concept under realistic operational conditions before proceeding to engineering development.

It is good to take a risk; however, when an implied commitment to production is involved, the technology should be operationally proof tested prior to commencing Full-Scale Development. On the other hand, avoid the temptation of thinking that anything is "state-of-the-art" until it is working in the field.

4. PROOF OF PERFORMANCE ON MAJOR CRITICAL ISSUES

In evaluating the initial DCP (or its equivalent), it is important to ensure that the tests to be conducted during the period from DSARC I to DSARC II address the major critical issues, especially those technological issues which are identified in the DCP.

Each test should have a single objective if possible, and the objective should be simply stated. A plan for the conduct of the test and the data collection, reduction, and analysis must be in sufficient detail so that one can readily evaluate the performance of the system whether or not the test objective can be met. A relationship between the identified performance parameters and the test results should be established prior to the conduct of the test. Further, the set of objectives for each of the tests should be clearly related to the program objective as defined in the DCP. When this relationship is not clear, amplifying data should be required.

The design of the set of tests to demonstrate feasibility prior to DSARC II should be based on a building block concept, with high technical risk items being tested early and with subsequent tests incorporating more and more of the hardware until the complete system concept has been demonstrated feasible.

Also, if any subsystem is being tested as a complete assembly, it should be examined to ensure that it is truly state-of-the-art and has been previously proven.

5. PROOF OF PERFORMANCE OF SOFTWARE

Test and evaluation should ensure that software products are tested appropriately as described in Chapter II and Annex A.

As previously stated, software has often been developed more as an add-on than as an integral part of the overall system. Software requirements need the same consideration as hardware requirements in the Validation Phases. Usual practices often do not sufficiently provide for testing the software subsystem concept. Often the facilities available to contractors for software development and verification are critical to schedule and cost.

6. PROOF OF PERFORMANCE OF HUMAN FACTORS CONCEPTS

<u>At an appropriate time in concept definition or Development Phase, T&E should authenticate the human factors concepts embodied in the proposed system design, examining questions of safety, comfort, appropriateness of man-machine interfaces, as well as the number and skill of the personnel required</u>.

The numbers of personnel required should be validated against both operational and maintenance requirements. Testing early versions in the "human acceptability and compatibility" environment is extremely important. This will also help to validate the manning requirements.

F. ISSUES: Operational Realism/General

The following are checklist items contained in this section:

1. Testing in Degraded Modes
2. Evaluation of Testing with Pre-Operational Equipment
3. Effect of Instrumentation on Test Realism
4. Joint Tests
5. Realism in Demonstrations
6. Realism of Maintenance and Repair in Testing
7. Operational Reliability Estimation in IOT&E
8. Effect of Observers on Test Realism
9. Justification for Realistic OT&E

1. TESTING IN DEGRADED MODES

<u>The system concept and possible implementations must not hinge on the requirement for the system or subsystems to be finely tuned when the expected operational environment suggests that this will not be likely.</u>

The system should degrade gracefully as a result of detuning caused from expected operational usage. If the capability to keep the system peaked is expected to degrade with operational use then tests should be conducted under the degraded conditions.

2. EVALUATION OF TESTING WITH PRE-OPERATIONAL EQUIPMENT

<u>Results of tests conducted during exploratory development and which most likely have been conducted on brassboard, breadboard, or modified existing hardware should be evaluated with special attention to items such as:</u>

(a) The packaging of the hardware may significantly affect the performance characteristics such that the suggested proof of validation is inconclusive.

(b) Scaling laws may invalidate the findings or introduce new technology problems.

(c) The laboratory type environment in which the hardware was tested may preclude the generation of data needed to validate that the concept and technology approach will be applicable to an operational environment.

(d) The tests may not include signals and noise sources representative of those that might be expected in an operational environment.

3. EFFECT OF INSTRUMENTATION ON TEST REALISM

<u>The constraints to be placed on the test because of the range and instrumentation are of prime importance</u>.

As previously stated, before DSARC II the test facilities and instrumentation requirements to conduct operational tests should be identified, along with a tentative schedule of test activities. The applicability of existing test ranges and the adequacy of current facilities and instrumentation should be verified. Insofar as possible, alternative approaches

 Preceding page blank

(different ranges, etc.) and instrumentation improvements needed should be specified. If range and instrumentation factors are found to cast significant doubt on the meaningfulness of the test data because of a lack of operational realism, the steps necessary to assure meaningful data should be identified and planned before DSARC II.

4. JOINT TESTS

Joint service operational test and evaluation should be considered for those weapon systems which require new operational concepts involving other services.

Emphasis in the joint tests should include investigation of the impact on the effectiveness of the weapon system of such aspects as CCC, target acquisition, damage assessment, and nominal types of countermeasures. If joint testing is recommended, an analysis of the impact of this type of demonstration on time and resources needed in the program and the additional resources needed to execute the joint tests should be conducted before DSARC I.

5. REALISM IN DEMONSTRATIONS

Demonstration and acceptance tests, as well as tests intended to evaluate performance under operational conditions, should always be conducted under conditions as close to those anticipated in practice as possible.

On the other hand, test conditions during development should be determined by the primary objectives of that test, rather than by more general considerations of realism, etc. Whenever a non-tactical, non-operational configuration is dictated by test requirements, the results of the tests should not be challenged by the fact that that configuration was not tactical or operational.

6. REALISM OF MAINTENANCE AND REPAIR IN TESTING

Prior to the decision to go into Full-Scale production of the weapon system, a complete technical/maintenance data package must be prepared and tested to ensure that the system can be maintained.

The testing of this package should be considered first as part of DT&E and then as part of the IOT&E of the system. Criteria for successful demonstration of this package should be established in both types of tests.

7. OPERATIONAL RELIABILITY ESTIMATION IN IOT&E

<u>IOT&E can provide valuable data on the operational reliability of weapon systems which cannot be obtained through DT&E</u>.

Factors such as operator error failures and apparent random failures should be looked for in the operational tests and investigated to determine if serious problems are underlying reasons for the failures. Especially important is the procedure used to evaluate the operational reliability of the system as determined by the relatively small amount of, but significant, data obtained through IOT&E and the large amounts of data on hardware design reliability collected through DT&E. Further, the maintenance practices should be carefully studied to assess their impact on the observed operational reliability obtained through IOT&E.

8. EFFECT OF OBSERVERS ON TEST REALISM

<u>Test conduct can be influenced by the actions of the observers and umpires</u>.

These people can provide important clues to the participants of operational suitability testing and in that way lessen the validity of the test. For example, in situations where air/ground duels are to be conducted, briefed observers who look in the direction of the aircraft, might inadvertently tip-off the direction of approach to the ground party in the duel. Similarly, concentrations of observers at a certain location may clue the aircrews where to search first for the ground targets.

9. JUSTIFICATION FOR REALISTIC OT&E

<u>Operational testing is essential, but it is also expensive and time consuming</u>.

Be sure in advance that the value received is worth its weight in not-delivered systems. Think in terms of:

(a) Involving operational groups in test planning and in establishing measures of effectiveness, so that the outcome of the tests will be accepted as being operationally significant.

(b) Determining whether the scope of the planned tests will provide sufficient data to justify any change at all in the eyes of potential users.

(c) Comparing the scope of proposed tests against checklists of issues frequently raised at major decision milestones, to assure that the data needed for such decisions will be forthcoming to the extent this is possible from testing alone.

(d) Recognizing in the formulation of test plans that major system decisions are judgments based on a wide range of qualitative considerations, rather than on statistical compilations, and that the outcome and limitations of operational tests must be comprehensive and meaningful to the decision makers as well as to the testing community.

F. ISSUES: Operational Realism/Personnel

The following are checklist items contained in this section:

1. Use of Appropriate Personnel During Test
2. Training Personnel for Tests
3. User Participation in Testing
4. Test Planning Personnel Qualifications
5. Continuity of OT&E Personnel in Test Planning
6. OT&E Pre-Test Training and Transition

1. USE OF APPROPRIATE PERSONNEL DURING TEST

Testers, evaluators and operators have quite different backgrounds and needs which affect the T&E of the weapon system.

Each has a different approach which has merit and utility at almost all points in the T&E program. A mix of these types is needed throughout the program. Early in the program, the lead emphasis should be from the tester, shifting to the evaluator and finally the operator, but at all times all parties and their needs should be coordinated.

2. TRAINING PERSONNEL FOR TESTS

Training plans and certification plans for test personnel should be established early in the Full-Scale Engineering Development Phase. Errors by test personnel are usually expensive and often cloud the reason for test failures.

3. USER PARTICIPATION IN TESTING

It is imperative that the Independent Test Agency participate in all of the T&E phases to ensure that the user needs are represented in the development of the system concept and hardware.

Initially, the Independent Test Agency should play an advisor role during the feasibility and engineering testing, and gradually take over leadership in the conduct of the testing program as it becomes more and more operational. This should facilitate the necessary communication and interaction between developing and user commands--especially needed during the DT&E and IOT&E phases.

4. TEST PLANNING PERSONNEL QUALIFICATIONS

The test director and/or key members of the test planning group within the project office should have significant T&E experience.

If the requisite experience does not exist at the appropriate levels within the project office, test plans may be based on too shallow or too naive a conception of the role and potential utility of the T&E process. All too often, key test personnel are assigned to T&E slots with little

prior exposure to T&E or its management, and with inadequately experienced support as well. The test planning group should have personnel experienced in engineering testing, development testing and operational testing. This experience should be available very early and all efforts should be made to encourage these people to remain with the weapon system project office through the T&E phases of the program.

5. CONTINUITY OF OT&E PERSONNEL IN TEST PLANNING

<u>The planners and evaluators for the OT&E of the production equipment can do a better job if they are initially involved in planning and conducting the IOT&E.</u>

The program plan should be reviewed to ensure that the FOT&E people are identified for IOT&E participation and that the personnel system of their service retains identity of these people for use in planning, conducting, and evaluating FOT&E which may not be run until a year or two afterwards.

6. OT&E PRE-TEST TRAINING AND TRANSITION

<u>In the initial conduct of OT&E, the participants should be given a period of time to dry run the scenario and to shake-down the instrumentation and the overall operation before key resources are expended in tests for record.</u>

In a properly planned OT&E program, the people will have completed proper individual training on the new system but the operational organization will not be able to conduct full unit training until the hardware, software, and support equipment are on hand. After the period when the unit is qualified as being operationally ready, it would be ready for assignment to OT&E testing.

F. ISSUES: Operational Realism/Threat and Environment

The following are checklist items contained in this section:

1. Offense/Defense Environment
2. Joint Service Operational Testing

1. OFFENSE/DEFENSE ENVIRONMENT

The OT&E plan should include offense/defense engagements in the environments in which the new system is expected to operate.

Offense/defense testing may be addressed in several phases, such as:

(a) One-on-one testing against existing U.S. counter systems and available simulators of the assumed threat.

(b) One-on-one testing against advanced U.S. technology which may be representative of a logical threat.

(c) Multiple vehicle testing in a multiple threat environment.

(d) Comparative testing of the new system with existing systems to estimate the increased capability.

Test range and resource requirements should be estimated, and, if inter-service testing is contemplated, preliminary plans for such testing should be coordinated with the cooperating service.

2. JOINT SERVICE OPERATIONAL TESTING

Joint service operational test and evaluation should be considered for those weapon systems which require new operational concepts involving other services.

As stated twice before, emphasis in the joint tests should include investigation of the impact on the effectiveness of the weapon system of such aspects as CCC, target acquisition, damage assessment, and nominal types of countermeasures. If joint testing is recommended, an analysis should be conducted before DSARC I of the impact of this type of demonstration on time and resources needed in the program and the additional resources needed to execute the joint tests.

G. REPORTING

The following are checklist items contained in this section:

1. Feedback of Test Results
2. Data Reporting Format
3. Data Collection on Subsystems and Components
4. Provision of Data for Modeling of Alternative Conditions and Scenarios

Preceding page blank

1. FEEDBACK OF TEST RESULTS

A good test program makes provisions for feedback of test results, during conduct of the testing, so as to influence:

(a) Course of the T&E program (test director, program manager).

(b) Trade-off decisions between modifying the system design and relaxing the operational requirements (program manager, operating/supporting commands, HQ).

(c) Missions, employment doctrine, tactics and constraints, tactical organization, etc. (operating command, operational units).

(d) Parts provisioning.

2. DATA REPORTING FORMAT

Establish a T&E reporting format for the program--insist on its use throughout the duration of the program.

Use this to:

(a) Establish a closed loop reporting and resolution process which assures that each test failure at every level is corrected by appropriate action, i.e., redesign, procurement, retest, etc.

(b) Establish a program-to-program crosstalk relative to T&E problems and approaches.

3. DATA COLLECTION ON SUBSYSTEMS AND COMPONENTS

When developing, testing and evaluating the various subsystems (and systems) of non-expendable weapon systems, each component of the systems should be numbered and a performance history kept which allows an analysis of that component's performance with respect to reliability, maintainability, availability, etc.

An analysis of failure modes should be made in advance so as to relate test results to the operational capability of the system when in a degraded condition.

4. PROVISION OF DATA FOR MODELING OF ALTERNATIVE CONDITIONS AND SCENARIOS

Develop techniques and system range instrumentation to provide the type of data in the proper form to allow economic, analytical, and mechanical simulation for alternate scenarios and combinations.

 Preceding page blank

Annex A

SOFTWARE

SOFTWARE

A. INTRODUCTION

This annex is intended to provide guidelines for program managers and program monitors in tracking the development of computer programs essential to the functioning of weapon systems. The purpose is to ensure that the software development is scheduled, performed, and tested with the same degree of attention to quality, schedule, and cost as is the hardware part of the system. It is assumed that the program manager's office will include in its staff experienced professionals who are skilled in programming design, implementation, testing, and support, and that the contractor or service developer will bring to bear the necessary talents for excellence in program design, implementation, testing, and support. It is also assumed that the program manager and his staff will provide sufficient information on overall system and software objectives to enable the developer to prepare two essential documents prior to development of test plans. These essential documents are the Program Functional Description and the Program Logic Description.

A number of checkpoints at which developer and program manager achieve agreement on critical issues are necessary to accomplishment of a successful development. The following list is a suggestion for the timing and critical issues to be covered at the specific points.

Checkpoint 1: Timing: At the start of the development.

Critical issues:

1. User/developer agreement in statement of and interpretation of requirement.
2. Establishment of the changes policy.
3. Establishment of the development plan.
4. Determine source of hardware required for software development.

Checkpoint 2: Timing: Upon completion of Program Functional Description.

Critical issues:

1. Reaffirm user statement of requirements.
2. Identification of potential problems in interfaces,

performance, diagnostics, human factors, standards compliance.

3. Adequacy of resources.
4. Development schedules.

Checkpoint 3: Timing: Upon Completion of Program Logic Description.

Critical issues:

1. Documentation of proposed data flow, logic flow, and program organization to implement each required function.
2. Determination of interfaces between segments of the program.
3. Reaffirm user statement of requirements.
4. Adequacy of computer facilities to accommodate program.

Checkpoint 4: Timing: Upon completion of test plan.

Critical issues:

1. Completeness and consistency of test plan with Program Functional Description and Program Logic Description.
2. User approval of test plan/criteria.
3. Credibility of schedule and cost planning.

Checkpoint 5: Timing: After critical functions in the program have been programmed.

Critical issues:

1. Verification in coordination with the user that critical functions have been completely and adequately covered by programming.

Checkpoint 6: Timing: After all testing is complete.

Critical issues:

1. Verification in coordination with user that the program meets all functional requirements.
2. Verification that program meets all specifications and user requirements.
3. User acceptance of test results.
4. Verification that program documentation is acceptable.
5. Verification that program support is feasible and plans for support are complete. (Support includes distribution, installation, training, publications, corrections to programs, updating of programs, development of field tests for user, etc.)

This list of checkpoints is intended only as a basic outline, leaving unsaid many details on procedures, working arrangements, record keeping, scheduling, etc. Likewise, it makes no attempt to provide guidance to programmers inasmuch as developers will have their own design methods and review procedures.

Testing, however, is the prime concern of the Task Force, and the following contains guidance on developing test plans. In what follows, not all items discussed may be applicable to every system. Furthermore, it is not comprehensive. Some systems, because of their nature, may have additional requirements that are not foreseen here. Implementation of the test plan assures that the system is satisfactorily tested.

Unit testing, a necessity in the testing of any system, is not presented in this document. This is the testing by a programmer of his code before incorporating it into the system. Procedures should be established to ensure that this testing is done exhaustively.

B. TEST PLAN OVERVIEW

The test plan must involve two major elements. The first is the design of the test cases. The system specifications form the basis for derivations of an exhaustive list of the functional variations. As the list is developed, test cases are designed to exercise each variation. A matrix of test cases, versus variations, provides a means of measuring the extent of coverage.

The second element of the approach is measurement. Unexecuted code (functions) must be detected and exposures evaluated. The test streams may then be expanded to cover those exposures.

Goals pertaining to the percentage of variations to be tested and the percentage of conditional branches executed should be established. They should be at a level which will assure the program manager that his software has been adequately tested. Completion of the testing effort would then be determined by achievement of the goals, not by schedules.

An integral part of the test plan must be detailed development and testing schedules. Each test plan must include a discussion of how the following areas will be tested:

- Reliability/Availability--the objective is to eliminate product incidents. This means that no software errors will result in reinitialization.
- Serviceability/Maintainability--provide for effective problem determination, problem diagnosis, and repair.
- Compatibility--the ability of a user to transfer from one program to another and continue to execute the jobs he has been executing.
- Usability--evaluate human factor characteristics.
- Capability--the ability of the program to function at various levels of stress.
- Security/Integrity--the ability of the program to protect data.
- Publications--the examples, limits, and externals specified in the publications are accurate and executable.

C. TEST PLAN CHECKLIST

Nature of Development Activity

Dependencies & Interfaces

Software
Hardware

Identification of Variations

Major Testing Areas

Function
Environmental Testing
Configuration Testing
Compatibility Testing
Limits Testing
Error Messages & Conditions
Publications Examples
Recovery Testing

Performance Testing

Stress & Load Testing

Additional Testing Considerations

Reliability/Availability
Serviceability/Maintainability
Usability
Security/Integrity

Test Criteria

Entry Criteria
Exit Test Cases
Exit Criteria

D. TEST PLAN OUTLINE

1. Nature of Development Activity

Give a brief abstract of the nature of the development activity and the approximate size of the effort in terms of the number of modules affected or the amount of code required.

Include copy of description or a reference thereto and Checkpoint Plan documentation pertinent to the test plan, e.g., development schedule. If these documents do not contain the names of new/changed modules, include the names here.

2. Dependencies and Interfaces

a. Hardware

- Identify any dependencies on hardware that are not available at the coder's location.
- Identify commitments to obtain this hardware.
- Identify any unique critical dependencies upon hardware that are available at the coder's location and contingency plans in the event of nonavailability.
- Identify any hardware standards to include communication interfaces, applicable to this development.

b. Software

- Identify all dependencies on software that are not available at the coder's location. Include dependencies on other products and on drivers.
- Identify all new interfaces with other parts of the product or include a copy of the specifications that contain these.
- Identify commitments to obtain required software in sufficient time to adequately test interactions before integration.
- Identify significant internal development checkpoints.
- Identify any software standards applicable to the development.
- Identify standard data elements and code applicable to this development.

3. Identification of Variations

Identification of all syntactical and semantic variations stated in the Programming Functional Specifications. These variations are all candidates for test situations.

4. Major Testing Areas

For each of the following areas (or others as applicable) indicate the extent of testing planned, the origin and format of the test cases, and the procedures and tools to be used in conducting the tests. Also indicate where test cases are planned to cover two or more areas with the same test cases. In the case of previously released products, plans for testing the new code in any area should incorporate the plans for testing maintenance changes for that product which are scheduled for the same time period.

a. Function Testing

Verification that the specified functions match the programmed functions. This encompasses the following areas of testing:

- Programming Function Specification Testing--verification that the explicit functional specifications have been correctly implemented. Error injection techniques are recommended, where applicable, rather than simulation techniques.

- Programming Logic Specification Testing--verification that the explicit logic specifications have been correctly implemented.
- Interference Testing--verification that all programmed functions have been fully specified.

b. Environmental Testing

Verification by means of both test cases and procedures that the system operates in a realistic environment (i.e., the way that it is intended for a user to use it). It should cover such areas as:

- Running at peak or near peak load conditions for a sustained period of time.
- Utilization of such hardware configurations as are available.
- Testing on a driver.

c. Configuration Testing

Verify that the program operates within the hardware and software systems that support it.

- Hardware Configuration
  - Should exercise the hardware-dependent code.
  - Should exercise the code on various hardware configurations to verify that there are no hidden hardware dependencies.
- Software Configuration

  Verification that the function is viable in the supported software environments, e.g., sequential scheduling, multiprocessing, multiprogramming, etc.

d. Compatibility Testing

Verify that the program is consistent with any other program(s) with which it claims compatibility. It should cover such areas as:

- Previous versions of the same program.
- Other design levels of the same program.

e. Limits Testing

Verify that the program limits are correctly stated. The program should be tested outside of the limit, at the limit, and within the limit. This testing should include:

- External Limits
  - Verification of capacity, i.e., the quantity of input permissible under various storage levels.
  - Verification of the quantitative constraints stated in the functional specifications, e.g., the size of a record, depth of nesting, number of characters in an identifier; e.g., design point.
- Internal Limits

  Verification of internal limits, e.g., table sizes, queue entries, etc.

f. Error Message and Error Condition Testing

Verify that the error handling facilities of the program operate as stated and that these facilities are sufficient for the errors that occur.

- Force every error message and verify the accuracy and clarity of each. If the same error message appears for more than one error or can appear at significantly different times in the execution of the product, then these situations should also be covered.
- Plans for introducing various error conditions, for example:
  - Operator errors
  - Source language errors
  - Hardware failures
- Verification of interfaces with error handling routines.
- Provide a list of all new/changes messages and completion codes.

g. <u>Publications Example Verification</u>

Verify the validity of publications, e.g., figures in the storage examples, and tables concerning function(s) appearing in program documentation.

- Program documentation verification should include such things as:
  - Sample programs
  - Sample procedures
  - Examples
- Provide a list of all new/changes publications.

h. <u>Recovery Testing</u> (if applicable)

Verify that the Recovery Specifications are met under all environments. This should include the following:

- Verify proper creation and maintenance of the Recovery Environment.
- Simply stated, this requirement is to ensure that the proper recovery routine gains control at the proper time. This may be affected by the following four factors, each of which must be verified:
  - Verify that the correct recovery type was established.
  - Verify that proper conventions are observed.
  - Verify that the required parameters are effective on the recovery routine exits.
  - Verify that all routines which make a recovery routine known cancel that recovery routine before returning to caller.

- Exercise recovery code for all error types.
- Exercise recovery code under all entry conditions.
- Exercise recovery code under all critical interface situations.

5. <u>Performance Testing</u>

Identify how performance specifications will be verified.

6. <u>Stress and Load Testing</u>

Identify to what extent the program will be run at peak or near peak load over an extended period of time.

7. <u>Additional Testing Considerations</u>

For each of the following areas that are applicable, include a discussion of how the topic will be tested:

a. <u>Reliability/Availability</u>

The objective is to eliminate program incidents. This means that no software errors will result in reinitialization.

b. <u>Serviceability/Maintainability</u>

Provide for effective problem determination, problem diagnosis, and repair.

c. <u>Security/Integrity</u>

The code must conform to the specification.

8. <u>Test Criteria</u>

Select criteria to be considered necessary for entry into the testing phase and sufficient for exit from the testing phase.

a. <u>Entry Criteria</u>

List what criteria must be met before this testing phase will begin.

b. <u>Exit Test Cases</u>

List any test cases that are <u>required</u> to be successful before exit.

c. <u>Exit Criteria</u>

List the criteria that have been selected as being required for exiting the test phase.

Annex B

THE TEST AND EVALUATION GAP

THE TEST AND EVALUATION GAP

Figure 1 illustrates some of the typical factors involved in the Test and Evaluation Gap problem. The chart shows key events and phases after the go-ahead for full-scale development, which occurs as a result of a favorable DSARC (II) decision. Typical R&D phasing is shown, where the first year or so is used in designing and building the initial test hardware. The subsystems then move into engineering tests, including R&D qualification tests. In the second and third year the system tests are conducted. Some Military Preliminary Evaluations (MPEs) occur early. IOT&E tests would be conducted after the R&D system demonstrated adequate adherence to the contract performance specifications.

If the IOT&E is reasonably successful and the service only then requests and obtains production authority for equipment to be used in OT&E, there will be a delay before production hardware is available because of the production tooling and production hardware lead time. To avoid the gap, depending upon the calendar time of the DSARC and the annual DoD budget submission and congressional defense, limited production funds would have to be defended a minimum of about 8 months prior to the major production decision. With less fortunate phasing, the budgeting lead time might be 4 to 6 months longer. Note on the Figure that this would require defense of the limited production program before the completion of the R&D system tests.

The limited production would normally be used by the first operational unit or the evaluation unit to do unit training and to work up to operational readiness for follow-on OT&E with production hardware. If there were no limited production the T&E gap would last for about 2 years, from the completion of IOT&E to the initiation of follow-on OT&E.

As illustrated in Figure 1, under GAP solution, if it were decided at the onset of the full-scale development that an additional phase of OT&E were to be pursued during what was formerly a gap period, then funds for gap filler test hardware and resources would have to be defended within about a year after the R&D go-ahead. The funds would have to be committed for long lead time items early enough so that the gap filler hardware, which would evolve from R&D to production configuration, would be available initially

 Preceding page blank

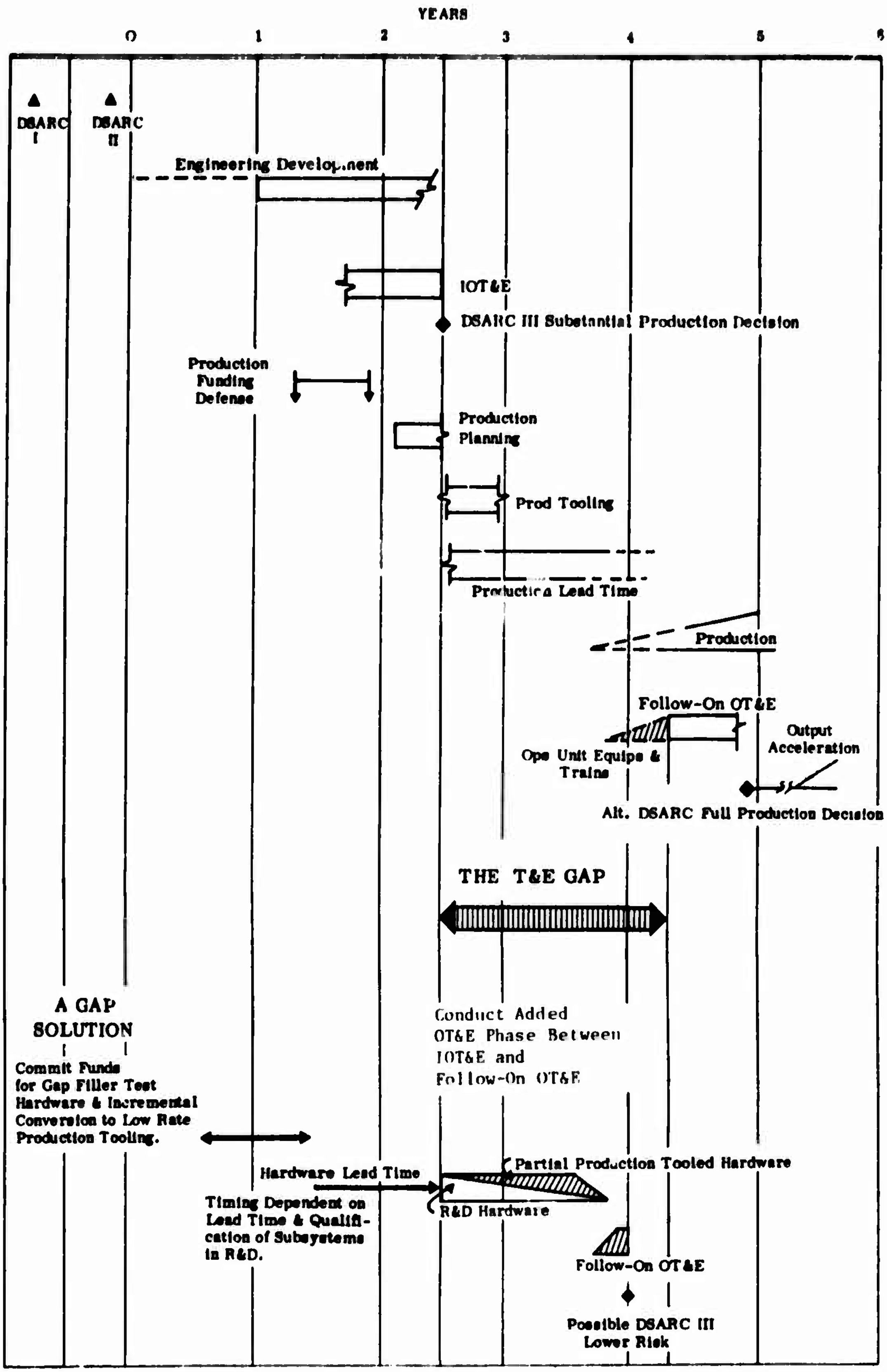

Figure 1. TYPICAL TEST AND EVALUATION GAP

at the end of the IOT&E. This additional OT&E phase would provide an additional year or two of operational experience before the major production output, thus providing a valuable opportunity to find and fix problems early, probably with R&D effort, hence minimizing costly modification programs which might be necessary if major production output followed a T&E gap. In addition, if the initial operations unit conducted this additional OT&E phase, the unit training would be accomplished and the unit should be ready to conduct follow-on OT&E as soon as the initial production hardware was available; hence, the initial operational capability (IOC) could be advanced several months. Certainly, the added years of experience during the former gap period should make the true capability at IOC much more effective.

It should be noted that the alternative of simply allowing the gap to exist, may be preferred when the effort to reduce the gap would require the commitment of a very large percentage (or amount) of the expected program cost before T&E assurance of a successful product could be obtained. Also, non-expendable system acquisition programs, such as aircraft developments, can continue to fly the R&D hardware during the gap period, but the stop and go in the building of aircraft is costly and key OT&E issues, such as reliability of production equipment, could not be addressed.

In summary, the T&E gap between IOT&E and follow-on OT&E is costly because inertia in the program is lost; government, contractor and subcontractor manpower are cut back and then in a short time built up again; valuable time is lost which could be used for perfecting and learning to use the system; faults not discovered early can be more costly to fix after production acceleration; and the true operational capability data is delayed. The problems in closing the gap are that funds for additional hardware must be defended before the R&D program will have shown much progress as an operating system, and more funds are required for the program prior to the major production decision.

Appendix A

TASK FORCE - TERMS OF REFERENCE

DIRECTOR OF DEFENSE RESEARCH AND ENGINEERING
WASHINGTON, D C 20301

MEMORANDUM FOR THE CHAIRMAN, DEFENSE SCIENCE BOARD

SUBJECT: Study of Past Procurement

I have asked Dr. Eugene Fubini to form a Task Force which will undertake a thorough analysis of a number of past system acquisition programs to enhance our understanding of the role which test and evaluation should have had in the identification of their problems and to make recommendations for the role of test and evaluation in future programs. I wish this Task Force to be established as a part of the Defense Science Board.

A copy of my letter to Dr. Fubini with the Terms of Reference for this study is attached. Lt.Gen. Alfred D. Starbird (Ret), Deputy Director (Test and Evaluation), ODDR&E, is the responsible deputy, and Mr. Howard Kreiner, Civilian Staff Assistant, Office of Assistant Director (Strategic and Support Systems Test and Evaluation) is the staff action officer for this Task Force.

John S. Foster, Jr.

Attachment
Ltr to Dr. Fubini

DIRECTOR OF DEFENSE RESEARCH AND ENGINEERING
WASHINGTON, D C 20301

14 Nov 1972

Dr. Eugene G. Fubini
Suite #816
1411 Jefferson-Davis Highway
Arlington, Virginia 22202

Dear Gene:

In the past few years there have been a number of reviews and studies of past and on-going weapons system acquisition programs, looking for means of avoiding or overcoming problems such as cost and schedule overrun, and system deficiencies in performance, reliability, and maintainability. Test and Evaluation activities have been looked at peripherally during some of these reviews, and some useful results have been obtained.

However, there has not been a major effort to investigate the possibility that effective testing could have resulted in earlier discovery and action on system problems.

I believe that a more complete investigation of representative programs would enable us better to understand how to improve our test and evaluation activities, where to concentrate more heavily and how to give our test and evaluation activities their highest potential payoff.

To conduct this investigation, I propose to establish a Task Force under your Chairmanship as a part of the Defense Science Board. I request that you assemble a select group to serve on the Task Force, to conduct the investigation of a group of specific programs. Please select the programs for study in coordination with Lt. Gen. A. D. Starbird (Ret.), my Deputy for Test and Evaluation. General Starbird will provide a full time staff member to your Task Force, and arrange for additional professional staff assistance through a contractor to be selected.

Your Task Force should conduct its investigations so as to establish for each program:

a. Whether the program had cost, schedule, or performance difficulties; from what specific aspects of the program these difficulties arose; and when the difficulty first became apparent (e.g., during design verification testing, acceptance testing, operational testing, or after deployment).

b. For each program and specific difficulty, was the discovery of the problem as early as reasonably could be expected? If not, what additional test measure reasonably could have been taken that might have found the difficulties? What test changes in the testing of future similar programs would appear warranted?

c. Based on the analysis of the entire group of programs, what areas and what potential problems should we examine more thoroughly and through what type and phase of testing? Further, are there areas in which excessive testing has been or is being carried out?

I expect that a year will be needed to address these questions. However, during this year we will work directly and closely with you in order to insure that the Task Force is working on the most important issues and that the Department is getting full benefit from early results of the Task Force's study.

Sincerely,

John S. Foster, Jr.

Enclosures
Memo for Chairman, DSB
Ltr for Prosp Task Force Mbr
& Distr List

Appendix B

DOD DIRECTIVE 5000.3 TEST AND EVALUATION

January 19, 1973
NUMBER 5000.3

DDR&E

Department of Defense Directive

SUBJECT Test and Evaluation

Refs.: (a) DoD Directive 5000.1, "Acquisition of Major Defense Systems," July 13, 1971
(b) DepSecDef multi-addressee memorandum, "Conduct of Operational Test and Evaluation," February 11, 1971 (hereby cancelled)
(c) DepSecDef multi-addressee memorandum on the subject of the role of DDR&E in test and evaluation as related to the DCP System, April 21, 1971 (hereby cancelled)
(d) DepSecDef multi-addressee memorandum, "Test and Evaluation in the System Acquisition Process," August 3, 1971 (hereby cancelled)

I. PURPOSE

This Directive establishes policy for the conduct of test and evaluation by the Military Departments and Defense Agencies (hereinafter referred to collectively as "DoD Components") in the acquisition of defense systems (Sections III through VI). In addition, it codifies the responsibilities of the Deputy Director of Defense Research and Engineering, Test and Evaluation (DD(T&E)), which were previously promulgated by references (b), (c), and (d)(Section VII).

II. CANCELLATIONS

References (b), (c), and (d) are hereby superseded and cancelled.

III. SCOPE AND APPLICABILITY

The provisions of this Directive encompass major programs of defense systems acquisition as designated by the Secretary of Defense (described in Section II., of reference (a)) and apply to all DoD Components that are responsible for such programs. In addition, it provides principles to be applied by the DoD Components in their acquisition of Defense Systems that do not fall in the "major acquisition programs" category.

IV. POLICIES AND PRINCIPLES

A. General.

1. Test and evaluation shall be commenced as early as possible and conducted throughout the system acquisition process as necessary to assist in progressively reducing acquisition risks and in assessing military worth.

2. Acquisition schedules will be based, inter alia, upon accomplishing test and evaluation milestones prior to the time that key decisions which would commit significant added resources are to be made.

3. Before the initiation of development of a new system, test and evaluation using existing systems, or modifications thereto, may be appropriate to help define the military need for the proposed new system and to estimate its military utility. Determination of military worth, need, and utility will be accomplished in accordance with other DoD directives.

4. All test and evaluation activities shall consider environmental issues and provide assessments for review as early as possible in the test planning cycle. (See DoD Directive 6050.1.)

B. Development Test and Evaluation (DT&E). DT&E is that test and evaluation conducted to: demonstrate that the engineering design and development process is complete; demonstrate that the design risks have been minimized; demonstrate that the system will meet specifications; and estimate the system's military utility when introduced. DT&E is planned, conducted, and monitored, by the developing agency of the DoD Component, and the results thereof are reported by that agency to the responsible Military Service Chief or Defense Agency Director.

1. DT&E shall be started as early in the development cycle as possible and include testing of component(s), subsystem(s), and prototype or preproduction model(s) of the entire system. Compatibility and interoperability with existing or planned equipments and systems shall be tested.

2. During the development phase following the Program Initiation Decision (Milestone I), adequate DT&E shall be accomplished to demonstrate that technical risks have been identified and that solutions are in hand.

3. During the Full-Scale Development phase and prior to the first major production decision, the DT&E accomplished

shall be adequate to insure: that engineering is reasonably complete; that all significant design problems (including compatibility, interoperability, reliability, maintainability, and logistical considerations) have been identified; and that solutions to the above problems are in hand.

4. For those systems which have a natural interface with equipment of another Component or may be acquired by two or more Components, joint DT&E may be required. Such joint testing will include participation and support by all affected Components as appropriate.

C. Operational Test and Evaluation (OT&E). OT&E is that test and evaluation conducted to estimate the prospective system's military utility, operational effectiveness, and operational suitability (including compatibility, interoperability, reliability, maintainability, and logistic and training requirements), and need for any modifications. In addition, OT&E provides information on organization, personnel requirements, doctrine, and tactics. Also it may provide data to support or verify material in operating instructions, publications, and handbooks. OT&E will be accomplished by operational and support personnel of the type and qualifications of those expected to use and maintain the system when deployed, and will be conducted in as realistic an operational environment as possible. OT&E will normally be conducted in phases, each keyed to an appropriate decision point. During Full-Scale Development OT&E will be accomplished to assist in evaluating operational effectiveness and suitability (including compatibility, interoperability, reliability, maintainability, and logistic and training requirements). OT&E will be continued as necessary during and after the production period to refine these estimates, to evaluate changes, and to reevaluate the system to insure that it continues to meet operational needs and retains its effectiveness in a new environment or against a new threat.

1. In each DoD Component there will be one major field agency (or a limited number of such major field agencies) separate and distinct from the developing/procuring command which will be responsible for OT&E and which will:

   a. Report the results of its independent test and evaluation directly to the Military Service Chief or Defense Agency Director.

   b. Recommend directly to its Military Service Chief or Defense Agency Director the accomplishment of adequate OT&E.

   c. Insure that the OT&E is effectively planned and conducted.

2. In addition, each DoD Component will provide within its immediate headquarters staff a full-time, strong, focal point organization to assist the independent OT&E field agency and to keep its Military Service Chief or Defense Agency Director fully informed as to needs and accomplishments.

3. Operational testing should be separate from development testing. However, development testing and early phases of operational testing may be combined where separation would cause delay involving unacceptable military risk, or would cause an unacceptable increase in the acquisition cost of the system. When combined testing is conducted, the necessary test conditions and test data required by both the DoD Component developing agency and OT&E agency must be realized. In addition, the separate Component OT&E agency must: insure that the combined test is so planned and executed as to provide the necessary operational test information; participate actively in the test; and provide separate evaluation of the resultant operational test information.

4. Acquisition programs will be so structured that at least an initial phase of operational test and evaluation (IOT&E) will be accomplished prior to the first major production decision adequate to provide a valid estimate of expected system operational effectiveness and suitability (including compatibility, interoperability, reliability, maintainability, and logistic and training requirements). Pilot production items will be employed for IOT&E wherever practicable. Prototypes, if they are reasonably representative of the expected production items, may be employed, where there otherwise would be delay involving unacceptable military risk or unacceptable increased acquisition costs.

5. For more complex systems, additional phases of OT&E may be required and performed with pilot or preproduction items subsequent to the first major production decision but prior to the availability of first production items. When production items are available in sufficient quantity, follow-on phases of OT&E adequate to meet the full objective outlined above will be accomplished by the appropriate DoD Component's independent OT&E agency.

6. For those systems which have a natural interface with equipment of another Component, or may be acquired by two or more Components, joint OT&E will be conducted where required. Such joint testing will include participation and support by all affected Components as appropriate.

D. <u>Test and Evaluation for Major Ships of a Class</u>. The long design, engineering, and construction period of a major ship will normally preclude completion of the lead ship and accomplishment of test

thereon prior to decision to proceed with follow ships. In lieu thereof, successive phases of DT&E and OT&E will be accomplished as early as practicable at test installations and on the lead ship so as to rapidly reduce risks and thereby minimize the need for modification to follow ships.

1. When combat system complexity warrants, there will be constructed a combat system test installation wherein the weapon, sensor, and information processing subsystems are integrated through their interfaces in the manner expected in the ship class. Adequate initial DT&E and OT&E of the integration of those subsystems will be accomplished thereon prior to the first major production decision on follow ships. To the degree practicable first generation subsystems will have been approved for service use prior to the initiation of integrated operational testing. Where subsystems cannot be service approved prior to the initial operational testing, their integration will be tested at the test site installation as early as possible in their acquisition cycle.

2. For new ship types incorporating major technical advancements not earlier proven in hull or non-nuclear propulsion design, a prototype incorporating these advancements will be employed. If the major technological advancements are contemplated in only some features of the hull or non-nuclear propulsion design, the test installation need incorporate only the applicable new features. Adequate test and evaluation on such prototype will be completed prior to the first major production decision on follow ships.

3. The prototyping of Navy nuclear propulsion plants will be accomplished in accordance with the methods in use by the Atomic Energy Commission. Construction of the lead and follow ships will be done in the sequence now being used.

4. For all new ship classes, continuing phases of OT&E on the lead ship will be conducted at sea as early in the acquisition process as possible for specified systems or equipments and, if required, full ship operational evaluation to the degree feasible.

5. A description of the subsystems to be included in any test site or test prototype, the schedules to accomplish test and evaluation, and any exceptions to the above policies will be set forth in the initial and any subsequent DCPs and approved by the Secretary of Defense.

E. <u>Test and Evaluation for One-of-a-Kind Systems</u>. For one-of-a-kind systems, or systems involving procurement of only a very few over an extended period, the principles of DT&E of component(s), subsystem(s)

and prototype or first production model(s) of the entire system will be applied. Compatibility and interoperability with existing or planned equipments will be tested. OT&E will be conducted as early as possible by the OT&E agency as necessary to provide a valid estimation of operational suitability and effectiveness.

F. Production Acceptance Test and Evaluation (PAT&E). PAT&E is test and evaluation of production items to demonstrate that the items procured fulfill the requirements and specifications of the procuring contract or agreements. It is the responsibility of each DoD Component to accomplish the necessary PAT&E throughout the production phase of the acquisition process.

G. Integrated T&E Plans. The DoD Component will prepare as early as possible in the acquisition process, and prior to initiation of Full-Scale Development, an overall test and evaluation plan to identify and integrate the effort and schedules of all T&E to be accomplished and to insure that all necessary T&E is accomplished prior to the key decision points. This plan will be kept current by the DoD Component.

H. Defense Systems Acquisition Review Council (DSARC)/Development Concept Paper (DCP) Procedures for Major Defense Systems.

1. The DCP prepared for use at the time of the Program Initiation Decision (Milestone I) for a major Defense System will identify the critical questions and areas of risk to be resolved by test and evaluation. It will also provide a summary statement of test objectives, schedules, and milestones. The DSARC in its review will determine the adequacy of the statement of questions and issues and of test objectives and schedules.

2. When the DoD Component proposes to initiate Full-Scale Development the revised DCP will give the results of T&E accomplished to that date, an updated statement of critical questions and areas of risk still needing test to resolve, and a detailed statement of test plans and milestones. The DSARC will assess and comment to the Secretary of Defense as to the adequacy of T&E progress and of planned T&E to occur prior to the first major production decision.

3. The DSARC in its review prior to the first major production decision will assess and comment to the Secretary of Defense as to the adequacy of test results to support a decision to proceed with major production and the adequacy of plans and schedules for any remaining testing.

4. In case of DCP revisions and DSARC reviews subsequent to the first major production decision, an updated assessment of test

results and plans and schedules for additional test and evaluation will be presented.

V. WAIVERS

A. In the case of major programs, any waiver of the accomplishment of the T&E as outlined in the approved DCP will be granted only by the Secretary of Defense.

B. For other than major programs, the DoD Components will designate the minimum threshold for definition of less than major programs. For such programs the waiver of the required T&E will:

1. Within the Military Departments, be granted only by the Secretary, the Under Secretary, or such Assistant Secretary as the Secretary may designate.

2. Within the Department of Defense Agencies, be granted only by the Director.

VI. EXCLUSIONS

Test and evaluation of nuclear weapons subsystems which are governed by other joint DoD/AEC agreements are excluded from the foregoing provisions of this directive.

VII. RESPONSIBILITIES OF THE DEPUTY DIRECTOR OF DEFENSE RESEARCH AND ENGINEERING, TEST AND EVALUATION (DD(T&E))

The DD(T&E) has across-the-board responsibility for OSD in test and evaluation matters. This responsibility includes:

A. Reviewing test and evaluation policy and procedures applicable to the Department of Defense as a whole and recommending changes he believes appropriate directly to the Secretary of Defense.

B. Monitoring closely the test and evaluation planned and conducted by the DoD Components for major acquisition programs and for such other programs as he believes necessary.

C. Assisting in the preparation of, and/or reviewing, the Test and Evaluation Sections of DCPs and Program Memoranda (PMs).

D. For major programs, reporting to the DSARC and the Worldwide Military Command and Control System Council as appropriate, and directly to the Secretary of Defense for such programs, at each major milestone decision point his assessment as to the adequacy of the identified critical issues and questions to be resolved by test and evaluation, test plans and schedules, and the adequacy of the accomplished T&E to justify the action recommended for that milestone decision.

E. Monitoring closely such joint testing as is accomplished by the DoD Components in connection with their planned acquisition of specific systems. In addition, initiating and coordinating the accomplishment of such additional joint testing as is necessary, with specific delegation to an appropriate Component (or Components) of all practical aspects of the joint test.

F. Coordinating and reviewing the test and evaluation of foreign systems for possible DoD use.

G. Fulfilling OSD responsibilities for the National and major Service test facilities.

H. Monitoring, only to the extent required to determine the applicability of results to weapon system acquisition or modification, that test and evaluation:

   1. Directed by the Joint Chiefs of Staff which relates to the Single Integrated Operational Plan (SIOP) operational factors.

   2. Conducted primarily for development or investigation of organizational or doctrinal concepts.

To accomplish these duties, statements of critical issues for DCPs/PMs, test plans for their resolution, and test results will be made available to DD(T&E) at his request as early as developed.

VIII. <u>REPORTING REQUIREMENTS</u>

The reporting requirements prescribed herein are exempt from formal approval and control in accordance with III.D.3., of DoD Directive 5000.19.

IX. <u>EFFECTIVE DATE AND IMPLEMENTATION</u>

This Directive is effective immediately. Each DoD Component which has authority and responsibilities under reference (a) will implement this Directive within 60 days and will forward three copies of each implementing document to the Director of Defense Research and Engineering.

Kenneth Rush
Deputy Secretary of Defense